THE ENTANGLED CYBERSPACE: AN INTEGRATED APPROACH FOR PREDICTING CYBER-ATTACKS

A thesis submitted for the degree of Doctor of Philosophy

By

Ruth Eneyi Ikwu

College of Physical Sciences, Department of Computer Science

Brunel University, London

September 2018

ABSTRACT

Significant studies in cyber defence analysis have predominantly revolved around a single linear analysis of information from a single source of evidence (The Network). These studies were limited in their ability to understand the dynamics of entanglements related to cyber-incidents. This research integrates evidence beyond the network in an attempt to understand and predict phases of the kill-chain across the information space.

This research provides a multi-dimensional phased analysis of the traditional kill-chain model using structural vector autoregressive models. In the 'Entangled Cyberspace Framework', each phase of the kill-chain corresponds to a single dimension of the information space based on time observations of certain events. Events are represented as time signals, where each phase is characterised by multiple time signals representing multiple events on that phase. Multiple time signals are analysed using structural models for multiple time series analysis (Vector Auto-Regressive models). At each phase of the kill-chain, we perform a lagged co-integration analysis of events across the information space. This nature of analysis detects hidden entanglements that characterise events in the kill-chain beyond the network. The measured prediction accuracy and error measured at each stage of the experiment represents the usefulness of selected events in characterising the defined stage of the kill-chain.

The entangled cyberspace, in theory, is the fusion of three conceptual foundations: a) A multi-dimensional characterisation of cyberspace, b) A sequential phased model for perpetrating cyber-attacks and c) A structural model for integrating and simultaneously analysing multiple sources of evidence. It starts with the characterisation of the information space into different dimensions of interest. The framework goes further to identify evidence sources across these characterised dimensions and integrates them in the analytical context under consideration (e.g. Malware Injection).

The concrete findings show that our approach and analytical methodology are capable of detecting entanglements when applied to a set of entangled activities across the information space. The findings also prove that activities beyond the network have significant effects on the nature of the unfolding cyber-attack vector. The predictive features of events across the kill-chain were also presented in this research as opinion and emotion drivers on the social dimension, packet data details and social and cultural events on the economic layer. Finally, co-integration detected between events across and within dimensions of the information space proves the existence of both inter-dimensional and intra-dimensional entanglements that affect the nature of events unfolding during the kill-chain (from the adversary's point of view).

The novelty of this research rests in the ability to hop across the information space for detecting evidential clues of activities that are related-to cyber-incidents. This research also expands the standard multi-dimensional information space to include SPEC factors as indicators of cyber-incidents. This research improves the current information security management model, specifically in the monitoring, analysis and detection phases. This research provides a methodology that accommodates a robust evidence base for understanding the attack surface. Practically, this research provides a basis for creating applications and tools for protecting critical national infrastructure by integrating data from social platforms, real-world political, cultural and economic events and the cyber-physical.

DEDICATION

To Hashem, Mom, Dad, Patrick, Jude, Justina and Chimdi, your support was profoundly appreciated. Thank You.

ACKNOWLEDGEMENTS

To the one who saw me through it all.

Thank you, Professor Panos Louvieris, for all your help, patience, understanding and supervision. Despite the rocky road, you were truly more than just a supervisor. I also want to acknowledge my second supervisor Zidong Wang.

To my parents Engr Edwin Ocha Ikwu and Mrs Benedicta Ikwu, I want to say a big thank you. For your moral, spiritual and financial support through this process. I could not have done this without you. And to my siblings, Patrick, Jude and Justina, this is for us.

I also want to thank the staff of the Computer Science Department who, after four years, seem like family. To Ela, Jeremy and Neela, thank you all.

To the very special friends who saw me through and helped in your way, Bibian Ogbuji, Chimdi Okolo, thank you all.

DECLARATIONS

The following papers have been published as a direct result of the research discussed in this thesis:
Paper 1: Ikwu Ruth, "Multi-dimensional structural data integration for proactive cyber-defence," *2017 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*, London, 2017, pp. 1-9.
Paper 2: Ikwu Ruth, Panos Louvieris "Natural Language Processing for Detecting Cyber-Related Discussions in Online Social Platforms. (In Review)

Contents

TABLE OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

1 CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

The state of current cyber defences at military, industrial and operational level have consistently proven inadequate in countering the ever-growing sophistication of the adversary and modern cyber-attacks. Powerful toolkits are readily and cheaply available to these attackers, and the networks are not prepared to handle these sophisticated attacks. Consequently, there is an increase in the application of proactive approaches to cyber situational awareness; such as measures to model patterns, groups and trends in security information (Doupé *et al.*, 2011; Skopik *et al.*, 2012; Leopold, 2015). In cyber analytical processes, it has long been the practice of studying the attacker's behaviour or pre-empting cyber-attacks by a critical analysis of cyber data (Wilson, Wiebe and Hoffman, 2005; Tartakovsky, Polunchenko and Sokolov, 2013). In addition to these techniques, mathematical and statistical models are being developed to model and understand network behaviour in various attack scenarios including correlated attack modelling for automating attack scenario recognition (George Mason University, 2010), Markov-chains for anomaly detection (Barford *et al.*, 2010), network anomaly detection with time series modelling of network packets (Münz, Li and Carle, 2007), Bayesian event *classification* (Kruegel *et al.*, 2003), clustering and classification models for botnet detection (Raghava, Sahgal and Chandna, 2012) and graph models for links in networks (Canright and Eng??-Monsen, 2008). For the best part of it, these techniques can help defenders in real-time cyber incident scenarios. However, most of these techniques are either modelled after-the-fact or based on attributes from some past cyber incident. Given the evolutionary rate of cyberspace and cyber activities, it is no longer adequate to base cyber defence protocols on these pre-assumptions but take a proactive disposition to the problem of cyber-attacks. Until recently (Hernández *et al.*, 2016), most cyber defence models offered a one-dimensional perspective to cyberspace as most of them were based on data from the network or the social layer.

Although these working models and techniques are continually re-defined by information security experts (Fischer and Keim, 2014; Abdullahi, Arif and Hassan, 2015), the sources of information feeding these models and techniques remain limited to a single dimension of cyberspace. The implications of these methods are 'after-the-fact' cyber awareness loops where a cyber-attack must have happened or at the least 'be happening' before learning a new signature.

This research argues that cyber-attacks are not independent of other events in the real world. Therefore, this research seeks to introduce a new perspective to analysing cyber-attacks based on a multidimensional structural approach to representing sources of information in cyberspace to build active predictive models for cyber-attacks. This research starts by characterising cyberspace as an integration of the physical and virtual realities of modern existence that consists of multiple smaller dimensions that make up the whole. This research presents the traditional cyber-attack kill-chain model from a multi-dimensional structural perspective by analysing the kill-chain across the various identified dimensions of cyberspace. In this thesis, we introduce these dimensions as sources of evidence for a multi-dimensional structural approach to pre-empting cyber-attacks. To achieve this, this research intends to capture the nature of entanglements between evidence on the various dimensions of cyberspace.

Cyberspace is the virtual link between geographically distributed cyber personas which creates the need for a virtual meeting point. Information is continuously generated between these cyber personas. Methods for addressing the problems in cyberspace must examine the entanglements of these cyber personas, their characteristics, activities and events across the multiple layers in which they exist in a time and space spectrum. Given this dynamic state of cyberspace, it is no longer logical to restrict the analysis of cyber behaviours to 'after-the-fact' methods. Pre-empting attacks against information assets and attacks enabled by cyberspace has become much more complex. Cyberspace provides unlimited access to malicious resources and a cost-effective platform to plan, recruit, train and execute attacks.

1.2 WHY STUDY THE PREDICTION OF CYBER-ATTACKS?

Over the last three decades, cybersecurity analysts have held a defensive approach to the offensive tactics of cyber adversaries. These techniques have kept cyber victims consistently one step behind their attackers. Also, the problem of the global commons (Lukasik, 2000, 2011) and the evolution of cyberspace ensures that the same information available to the defenders are also available to the offenders which consequently creates an unending loop of attack → defend → relearn → attack. Reactive approaches to cyber defence typically utilise firewalls, anti-viruses, anti-spyware, and virus and exploit detection programs to counter, for the most parts, known signatures of cyber-attacks that have happened in some past time. In most cases, some of the damage impacts the defender's network before mitigation strategies are put in place. If proper evidence represents events in cyberspace, there is a potential for modelling as an integration of webbed data, in which physical, social and virtual perceptions, interactions and relationships are linked in a time and space spectrum.

1.3 WHY USE ANALYTICAL AND STATISTICAL APPROACHES?

The case for analysis and statistics in cybersecurity is considerable covered in modern research (Lau, Xia and Li, 2012; Bar *et al.*, 2016; Hernández *et al.*, 2016; Shekhar, 2016). In the past, defenders invest heavily in threat prevention and identification (Forum and Reserved, 2011) as a reactive approach to cyber defence. This approach focuses on disrupting the cyber-attack kill chain (Engel, 2014) at the reconnaissance or delivery phase (Yadav and Rao, 2015). This approach is adequate for mitigating singular component cyber threats (Wlodarczyk and Hacker, 2014) such as email spams, phishing attacks and malware attacks. However, they may not be sufficient for mitigating multi-stepped, multi-dimensional attacks such as DDOS, APTs and similar advanced attacks.

Staying one-step ahead requires an analytical approach to cyber defence. Defenders must, therefore, supplement cyber-attack prevention with a thorough strategy for pending threat detection by understanding the relationships that exist between entities operating in cyberspace.

1.4 PROBLEM DEFINITION

The proliferation of cyber-attacks has been proven to be entangled with various other events in cyberspace (Gandhi *et al.*, 2011a). The entanglement of information in cyberspace provides a collection of webbed data that are linked to cyber-incidents and activities surrounding them. Research in cyber-situational awareness covers little on integrating multiple sources of evidence from uniquely identified strata of cyberspace to inform predictive models and proactive defence strategies. Given multiple sources of evidence from an entanglement of webbed data, we hope to develop techniques, methods and theories that best capture the nature of entanglements to predict cyber-incidents.

1.5 RESEARCH QUESTION

To date, frameworks and approaches that support cyber-incident prediction with such theoretical and practical synthesis of data across the various levels of the information space has been lacking in industry and research. The question of the entanglements that exist between events in the cyber-physical-social realm has not been investigated in research. In light of the above stated gaps,the researchers questions of interest are as follows:

- How can cyberspace be characterized to represent the various roles it plays in our daily lives. This characterization would help identify various streams through which data is generated?
- What are the various types and sources of data that exist in cyberspace that are predictive of cyber events?
- What is the nature of entanglements that exists between events across these various characterizations of cyberspace?
- What models, frameworks or approaches can capture these entanglements ad exploit them for active prediction of events in cyberspace?

1.6 RESEARCH AIM

Develop an approach and theoretical framework for predicting cyber-attack threats in the entangled and complex Cyberspace by characterising and extracting relevant data related to cyber-incidents across the layers of cyberspace to feed statistical and mathematical models for enhancing predictive accuracy. The technical aim involves identifying weak and co-integrated signals that reduce the predictive error of subsequent stages of the attack kill-chain.

1.7 RESEARCH OBJECTIVES

a) *Review the methods and algorithms for an integrated approach to pre-empting cyber-attacks.*

b) *To determine the features that defines the nature of entanglements in cyberspace.*

c) *Develop a multi-dimensional structural framework for integrating multiple sources of evidence in cyberspace to pre-empt cyber-attacks.*

d) *Assess the implications of the research findings and an integrated approach for pre-empting cyber-attacks in practice.*

1.8 RESEARCH SIGNIFICANCE

The dynamics of cyberspace provides opportunities to conduct complex cyber-related operations in multiple complex domains for both attackers and defenders (Ben-Asher and Gonzalez, 2015). Cybersecurity defences must, therefore, be sensitive to spot emerging cyber-attack vectors, real-time anomalies in cyber-security data and signs of under siege attacks. To this effect, defenders must continually analyse the vast amount of data with predictive techniques to stay one step ahead. Effective cyber analytical technics would create models that properly describe the dynamic relationships and nature of entanglements between components across layers of cyberspace. Consequently, this generates a data and context-driven approach to cyber situational awareness that captures the behaviours of all cyber entities existing in a specific domain.

The significance of this research rests in the ability to actively pre-empt a cyber-enabled incident with an approach for the untangling, identification, integration, representation and analyses of information from various levels of cyberspace for a proactive cyber situational awareness model. The possession

of an approach that can actively track and decipher causal links between events in cyberspace would put the defenders ahead of the attackers and enable proper cybersecurity operations.

1.9 RESEARCH APPROACH

This research is broadly divided into four main activities.

a) *The development of a theoretical framework and Hypothesis that forms the foundation for an integrated analytical approach.*

b) *Determine which measurement data in cyberspace are required to feed statistical models.*

c) *Investigate and create new algorithms for combining selected mathematical, statistical and computational techniques that can reveal early warning signs in complex information space.*

d) *Conduct practical and theoretical evaluations of methods.*

A theoretical framework for the analysis of cyber security data for cyber-attack prediction identifies the actors, events and techniques in the context of developing a cyber-security solution. Algorithms and methods previously applied are explored, compared and filtered down to a selected few based on a practical evaluation of performance. This is achieved by intensive qualitative research and a practical understanding of literature and techniques in the subject area.

The experiment used in this research is based on benchmarked data from a hypothetical scenario of 3 events. The scenario is designed to model the entanglement of activities within these three events and the entangled cyberspace framework is intended to pick up evidences of relationships between these events using appropriate methods. The scenario models the progression of activities across the various dimensions of cyberspace to the occurrence of a real-world cyber incident.

The quantitative research performed on a series of formulated hypotheses based on selected prediction models and re-evaluated in context. The formulation of hypotheses is based on benchmark datasets in cyber security analysis and test questions related to possible correlation, co-integration and causation between identified variables. These hypotheses are tested, and the resulting models are applied to the prediction of a real-world cyber-attack incident.

1.10 RESEARCH HYPOTHESIS

Following the stated objectives of this research, the researcher puts forward the hypothesis below as the investigative focus of this study. Structurally integrated data streams from multiple dimensions of cyberspace using the entangled cyberspace framework are predictive of the cyber-attacks. The main hypothesis is further broken down into multiple hypotheses to meet the aims and objectives of this research as stated in section 1.7 of this report.

1.10.1 Research Hypothesis 1

The phases in the perpetration of a cyber-attack can be characterised across the various dimensions of cyberspace.

1.10.2 Research Hypothesis 2

Given a multi-dimensional cyberspace, there contains a set of multi-dimensional features across various layers of cyberspace that are predictive of cyber-attacks.

1.10.3 Research Hypothesis 3

The entangled cyberspace model proposed in this research is capable of identifying predictive features of a cyber-incident across the various dimensions of cyberspace.

1.10.4 Research Hypothesis 4

The entangled cyberspace model proposed in this research is capable of predicting cyber-incidents in cyberspace.

1.11 THE LAYOUT OF THE THESIS

Chapter 2 Literature Review: of this study explores literature in the areas of cyber security, situational awareness and predictive analytics with time-based techniques to develop and design a theoretical and hypothesised framework for active prediction of cyber-attacks. This is achieved by providing a brief background on the current state of cybersecurity, an understanding of the current state of cyber-attacks and the existing defences to counter cyber-attacks. It goes further to characterise cyberspace to define the usefulness of data generated in cyberspace to the prediction of cyber-attacks.

Additionally, chapter 2 develops the theoretical and conceptual framework for the prediction of cyber-attacks in cyberspace. It further presents multiple hypotheses for the experimental design. The framework is put in the context of the cyber-attack kill chain and therefore offers a phased approach to the cyber-attack prediction problem. The Cyber-Physical-Socio-Economic space is presented as a model for characterising cyberspace as an entanglement of webbed data. Finally, it shows selected models in time series literature that are useful for analysing the structural data from cyberspace.

Chapter 3 Hypothesized Model and Approach and the Theoretical Development of the Entangled Cyberspace: develops the theoretical and conceptual foundations of this research. It critically discusses how the entangled cyberspace framework is developed using a combination of existing theories addressed in the literature. Thus, addressing how the current theories/models are used to ground the entangled cyberspace framework

Chapter 4 The Research Methodology: reviews contemporary research methodologies to establish a combination of them useful to this research. This chapter also presents the strategy for data collection and the methods of analysis used. In this chapter, the techniques and methods used in this research are presented and critically discussed. Specifically, vector autoregressive models are presented as a model for verifying the hypothesised theoretical framework, which is then extended into the cyber testing environment.

Chapter 5 Data Analysis: presents the analysis of data for the research experiments and begins the experimental design for the test of the multiple hypotheses developed in chapter three. The experimental design builds on suitable research methodologies identifies in the previous chapter. It begins by identification of variables useful for pre-empting cyber-attacks. It then describes methods and techniques for fusing data from various layers of cyberspace. Finally, it tests the usefulness of these variables to active cyber situational awareness using vector autoregressive models. Evaluation of these techniques is also presented to test the impact of cyber-attack prediction on active cyber-situational awareness.

Chapter 6 Discussion of Findings: critically discusses the finding identified in chapters two, four and five. This chapter also justifies the use of structural VARs in proactive cyber defence strategies to enhance cyber situational awareness. The research's contribution to theory is a robust framework for understanding the entanglements of events in cyberspace, thus improving the quality of cyber situational awareness strategies. In practice, this research provides active techniques for identifying features useful to the prediction of cyber-attacks and a method for analysing these features.

Chapter 7 Conclusion: rounds up the research process, evaluates the accomplishments and limitations of this research. It goes further to identify existing loopholes and recommends areas for further research in an attempt to close the research gap.

Figure 1-1 visualizes the structure of this thesis and steps implemented in each chapter to achieve the research aims and objectives. (see overleaf)

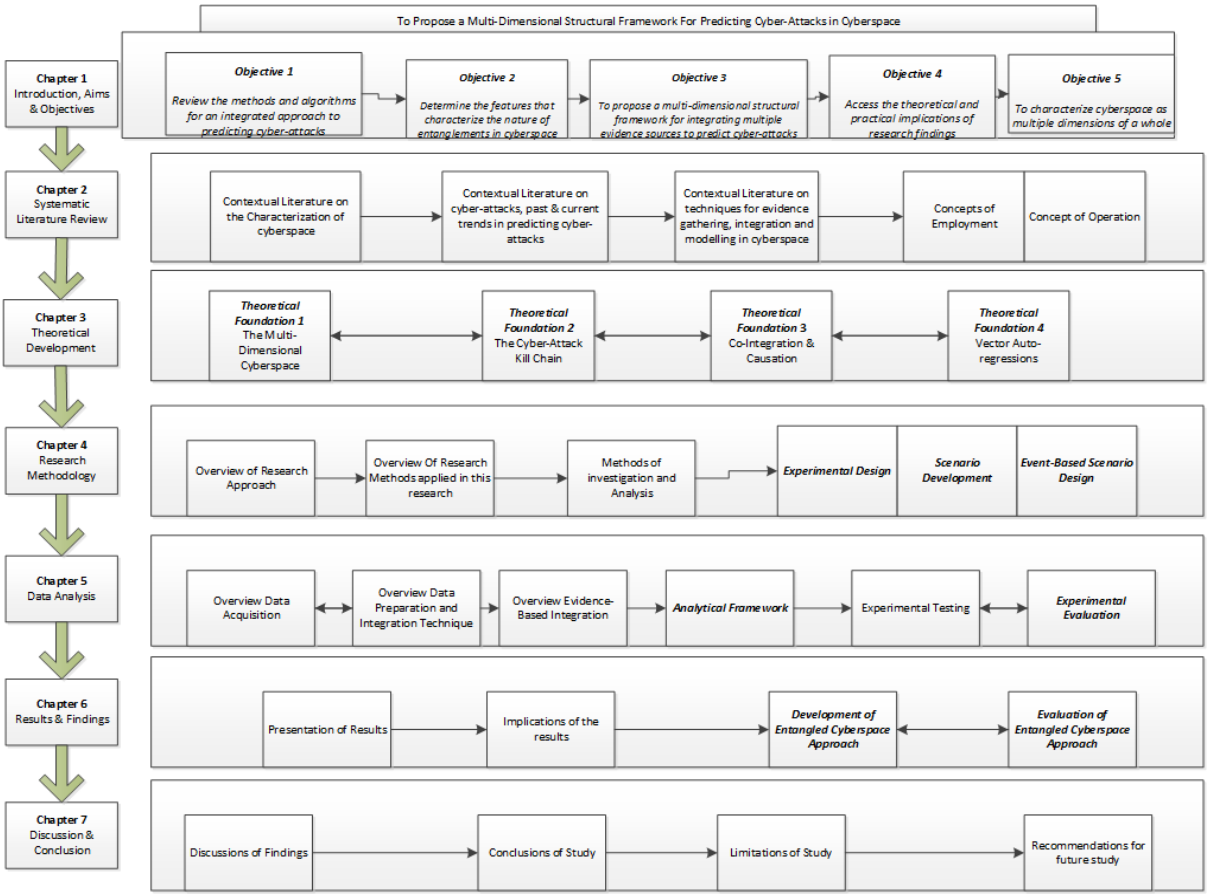

Figure 1-1: Structure of Thesis

2 CHAPTER 2: LITERATURE REVIEW

2.1 INTRODUCTION

This chapter explores the current literature in cybersecurity and information analysis in cyberspace. It begins with an introduction to cyberspace and the current landscape of cybersecurity. It further presents the effects of the new cyberspace individuals and modern existence. It also introduces a multi-dimensional characterisation of cyberspace and explores further the entanglements of information as it exists at various layers of cyberspace.

This chapter, as a critical analysis of information analysis in cyberspace, explores the current methods and techniques for analysing information in cyberspace and the current applications of statistical and mathematical models to areas of cybersecurity. Additionally, this chapter investigates the usability and evaluations of these methods in the context to which they are applied. This chapter further highlights data sources for information gathering as a pre-step to models' specification methods for data identification, extraction, transformation and representation.

Finally, the representation and analysis of multiple data vectors in multi-dimensional structural models are briefly introduced as a prelude to research methodology. These methods incorporate the techniques of multiple series analysis, change-point detection, co-integration and granger's causality, vector autoregressive models and structural equation models for the analysis of structural information.

In conclusion, this chapter delivers a critical analysis of existing frameworks for cyber situational awareness and builds up the foundation for the construction of a structural framework for combining methods in building predictive models for cyber-attacks.

2.2 SEARCH STRATEGY

This chapter explores current literature in cybersecurity, cyber situational awareness, cyberspace characterisation and techniques for analysing time-dependent data. The literature review attempts to identify and synthesise all empirical evidence that meets set criteria to fulfil the aims and objectives of this research. The search themes include 'cybersecurity and defence', 'cyberspace characterization', 'analytics in cybersecurity', 'co-integration', 'causation' and 'autoregressive models'. The search themes were passed through various academic databases including ACM Digital Library, IEEE Xplore, ScienceDirect, Wiley Online Library and SpringerLink. Generalised search engines used also include Google Scholar, Scopus and Microsoft Academic.

The search theme 'Cyber Security and Defence' was filtered to include only research published after 1999. This strategy keeps the results returned from the search relevant to the current state of cybersecurity and defence over the last two decades. Furthermore, the researcher restricts results from Co-integration and Causation search to those after 1999.

2.3 STRATEGY FOR CONDUCTING THE LITERATURE REVIEW

Based on the research questions and objectives, 3 major areas of investigation were identified; characterizing cyberspace, identifying features that characterize events across the multi-dimensional cyberspace and predicting cyber-incidents. To capture relevant research around these topics, the researcher conducts a literature review to identify gaps in the areas of interest and investigate how past studies have implemented the techniques used in meeting the research objectives. The researcher

investigate literature on characterizing cyberspace to develop a multi-dimensional representation of cyberspace based on the characterizing features of each layer. Additionally, the researcher reviews methods in time series analysis for predicting events recorded at equal space in time and identifies various model development and evaluation methods used in current research. The researcher also reviews other methods implemented in the various topic areas employed in addressing similar problems in the research space. The literature review is based on both legacy and current methods that address the problem area of interest.

2.4 CYBERSPACE IN MODERN DAY

Cyberspace has been defined by experts (Clark, 2010; Klimburg and Mirtl, 2012; United States Defense Force, 2013) as an artificial world where humans navigate in an information-based space. Critically, (MOD, 2016) recognises the importance of the interdependent nature of systems, people and networks in an operating space called cyberspace. Since the mid-1980s, cybersecurity experts have been observing an increasing gap between the evolution in the sophistication of the 'attacker' and the evolution of defences to counter these attacks. Given the constant development in contemporary modern technology, the increase in the number of interconnected devices in cyberspace has also increased. These discoveries have also led to an increase in attacker technique and sophistication. Unfortunately, while the attackers improve in their skills and attack execution tactics, the methods for countering these attacks have not experienced equivalent increase.

There is no standard definition of the term *'cyber'*, but the term is mostly used as a prefix for ideas related to internet networks, information technology and virtual realities *(COED)*. Information security experts (United States: US Army, 2010; United States Defense Force, 2013; Barnett, Smith and Whittington, 2014) mostly view the cyberspace as a network of internet and digital entities that support information flow. An interesting term cited in (Ning *et al.*, 2016) refers to cyberspace as a virtual and physical abstraction with all the information resources that create the interconnections among cyber entities. As a dynamic, interdependent environment, which is geographically less constrained than other environments, access is possible through virtual applications running on digital devices that store, process and transmit data (Li *et al.*, 2011).

In the past decade, cyberspace has gradually replaced the physical reality, and these digital communication devices have been deeply embedded in human lives. The role of cyberspace has thus progressively evolved over the last two decades from being a simple enabler of daily live processes to a critical necessity for modern existence. Shopping, financial transactions, communication, crime, education, entertainment are traditional processes of everyday lives that are conveniently replaced by digital processes in cyberspace. Consequently, for all these processes being dependent on digital communications and networks, information is also continually being generated in time.

From a cybersecurity perspective, at its core, the problem lies in a self-enabling cyberspace where all information to protect and harm information assets is provided by cyberspace itself. It is possible to access information to design a sophisticated cyber-attack (Olsen, 2013; Yadav and Rao, 2015). How these attacks may vary would be based on the skills and experience of the attacker. As a result, physical crimes and unlawful acts have found a way to exist in cyberspace anonymously. Terrorist groups, cybercriminals, government activists use cyberspace as an enabler for activities. Hathaway et al (Hathaway *et al.*, 2012) provides logical reasoning to the problem. The traditional rules guiding community-based interactions are based on the assumption of some sort of physical presence of one or more of the entity (ies) in question. These rules have slowly become obsolete in cyberspace. Unlike our physical world where physical interactions occur during day-to-day processes, the need for a physical presence is simply cut off in cyberspace (Zetter, 2014).

Cyberspace encompasses the people interacting, the devices connected and the information flowing within it. In recent times, the role of cyberspace has gradually evolved as an enabler of day-to-day lives. Cyberspace is a global village of inter-connected entities interacting through time. An interesting term cited in (Ning *et al.,* 2016) refers to cyberspace as a virtual and physical abstraction with all the information resources that create the interconnections among cyber entities.

2.5 FROM PHYSICAL TO VIRTUAL REALITY

The physical reality is the state of things as they exist in the real world. Over the last ten years, cyberspace has gradually replaced the physical reality, and these digital communication devices have been deeply embedded in human lives (Ning *et al.,* 2015). The role of cyberspace has thus progressively evolved over the last two decades from being a simple enabler of daily live processes to a critical necessity for modern existence. Shopping, financial transactions, communication, crime, education, entertainment are traditional processes of everyday lives that are conveniently replaced by digital processes in cyberspace. Consequently, for all these processes being dependent on digital communications and networks, information is also continually being generated in time.
Cyberspace is the virtual link between geographically distributed cyber personas who create the need for a virtual meeting point. Information is continuously being generated between these cyber personas. Methods for addressing the problems in cyberspace must examine the entanglements of these cyber personas, their characteristics, activities and events across the multiple layers in which they exist in a time and space spectrum. Given this dynamic state of cyberspace, it is no longer logical to restrict the analysis of cyber behaviours to 'after-the-fact' methods. Pre-empting attacks against information assets and attacks enabled by cyberspace has become much more complex.

2.6 CHARACTERISING CYBERSPACE

As technology evolved over the years, so has the dependence of modern society on networked infrastructure for communication and connectivity. In a simplistic, naive definition, cyberspace is a collection of multiple computing devices connected by networks for storing, transmitting and utilising electronic information. An alternative way to understand the nature of cyberspace is to demystify its purpose or the role it plays in human existence. (Clark, 2010) summarises the three main objectives of cyberspace:

i. Manipulation and exploitation of information.
ii. Facilitation and augmentation of communication among people.
iii. The interaction of people and information.

From these encompassing definitions, we can extract some common uses of cyberspace as relevant to modern existence: connectivity, communication and beliefs.

2.6.1 Cyberspace as a medium for inter-connectivity (The physical Space)

The physical infrastructure and logical building blocks of cyberspace support capabilities for connectivity between linked devices. The nature and rules upon which cyberspace exist are defined on the logical layer. For the physical infrastructure, it is important to note the main components that support connectivity sensors, actuators and networks. Sensors capture physical data and are low-cost devices with limited storage and computational capacities. Actuators, automate, control and convert the collected physical data into action commands. Intelligent computer networks support the linking

of digital devices and transmission of physical data through cyberspace. The physical objects establish connections and relations with other cyber entities in both real and digital worlds based on rules defined on the logical layer. Moreover, physical objects are owned and operated by human beings and therefore reflect social attributes that are in direct or indirect correlation with the affiliated persona.

2.6.2 Cyberspace as a medium for communication (Social Space)

Real-world personas control cyber-personas in cyberspace. The persona and cyber-persona are sometimes referred to as the offline and online modes respectively, of human beings in cyberspace (Ning *et al.*, 2016). The logical infrastructure of cyberspace provides capabilities for interaction and communication between multiple cyber-personas which in turn creates greater capabilities for sharing and collaboration in cyberspace. Sharing refers to an integration of social attributes and social events, exchange of thoughts and the creation of social inter-/intra relationships owned by real-world personas. Collaboration refers to a coming together of cyber-personas in real-time mainly to create intellectual content. This capability relies on communication infrastructure to support human learning and co-existence. Therefore, cyberspace as a medium for communication creates a social space of intelligent networks with streams of thoughts, perceptions, social affiliation and social relationships.

2.6.3 Cyberspace as an Enabler of beliefs

Given the social capabilities provided by the infrastructure of cyberspace, over the past three decades, cyberspace has also gradually become an influencer of beliefs. Real-world persons are affiliated with various social, political, economic and cultural ideas which shape ideologies. Cyberspace provides a medium for sharing ideologies and creating online communities based on ideological differences. Agents of ideological formation embedded in online social communities propagate ideas, notions and positions that support certain ideological perspectives.

Finally, (Ning *et al.*, 2016) summarises these functions of cyberspace in the formation of a cyber-human hyper-thinking space consisting of data, information, knowledge and thoughts. Data collected by sensors, being exchanged during communication sessions from ubiquitous devices are expected to create context-aware information that can be further mined to generate useful knowledge. The convergence of connectivity, communication, interaction and knowledge enabled by human personas encompasses the ultimate function of cyberspace in modern existence.

2.7 THE INTER-CONNECTED WORLD

The theory of inter-connectivity in cyberspace holds that people, things, entities, processes are inter-connected by the collaboration of the human, social and digital world (Ning *et al.*, 2016). The cyber-physical-social space is an interpretation of the theory of inter-connectivity in cyberspace in which physical perceptions, cyber interactions and social correlations are inter-connected through ubiquitous virtual reality. The underlying basic assumption is some form of inter-dependence between activities in the virtual reality, activities in the physical reality and activities in the social reality which is based on a perfect integration of the cyber-physical-social space with human, social and physical interactions.

Cyberspace, artificial and virtual realities are a generalised form of digital abstractions that support inter-connectivity and interactions between cyber entities. Cyberspace is assumed to be independent

of space and time constraints as its existence is a timeless virtual abstraction of the physical reality (Gotved, 2006). For example, the death of a person does not necessarily imply the death of associated cyber accounts. Access to cyberspace is possible via digital devices on independent networks. The cyberspace ecosystem also provides infrastructure for massive information management services, resource management and service management through public interfaces mostly running on distributed databases. This ecosystem is also supported by uniform standards and protocols to enhance the flexibility of usage, ease of access and create a self-evolving computing eco-system. Although cyberspace operates as a global domain consisting of independent sub-networks of technological infrastructure and data that potentially connects all interacting entities, the physical barriers of the real-world dimension (territory, land) are obsolete. In this way, it can provide the maximum number of people with the means of access, production and creativity within its new information society.

The physical space refers to real-world entities that interact directly or indirectly with cyber entities. The idea of physical space saturates modern existence as a concept for understanding the world around us and how we interact with its comprising entities. Thus, it encompasses all physical entities where physical interactivity is possible, heterogeneous interfaces, physical infrastructures and interactive environment required for seamlessly browsing through cyberspace. These include; all real-world components and the physical infrastructure that support the existence of independent networks. The facets of the physical space: Location, Distance, Size and Route (Hornecker and Buur, 2006) attribute physical entities and their mappings in cyberspace. A physical object can be mapped into cyberspace using geographical attributes (IP Addresses, Longitude, Latitude) so that its location replicates where the thing exist in the physical dimension in relation to other physical entities in the same space. The amount of time required to exchange information between two physical objects is a function of the distance between the two objects, the size of data being transmitted and the channel of communication (Feria, 2010).

The social space is an integration of social attributes, inter-personal relationships controlled by physical people and other physical objects. A physical entity creates direct and indirect relationships with other physical entities or cyber entities, therefore, establishing an eco-system of semantic relationships that mirror human social behaviour (Ning *et al.*, 2016). Physical people map characteristics or attributes of their personas into the social space as cyber personas. The personas and cyber-personas characterise human activities, social events, behavioural conventions, political administrations, public services for human social participation in cyberspace. A single physical user may be associated with multiple cyber personas however each cyber persona can be associated with one individual. Social interconnectedness enables the creation of online communities characterised by the active social presence, social participation, relationship creation, and collaboration and belief affiliations. These online communities sometimes mirror traditional real-world communities where members share and express similar political, cultural, social and economic ideologies. These three subdomains converge to a cyber-physical-social ecosystem held together by four main features; interconnectivity (Granovetter, 1982), interaction (Fu, 2016), integration (Maurizio, 2002) and intelligence (Freitas, 2013). These are considered to be enablers of the cyber-physical-social eco-system.

Interactivity refers to a two-way effect created when two or more entities meet each other. Entities in cyberspace can establish interactivity with other entities within their cyberspace dimension or with entities in other dimensions. For example; a web application accepts an input from a human, or a mobile phone establishes a one-touch payment process. The ease of access to cyberspace and robust infrastructure supporting distributed networks creates the platform for easy and speedy multiple interactive sessions between these cyber entities.

Inter-connectivity refers to a state of being connected or the linking of two or more entities in a given space (Fu, 2016). It is made up of multiple simple interactive processes happening simultaneously. Although this concept is often used when referring to the linking of two digital devices in an internet network, interconnectivity in cyberspace broadly indicates that physical objects, people, cyber personas and social dynamics establish seamless communication or relationship links between and within each other. This implies that entities may be linked via physical connections or social relationships and interactions across the various dimensions of cyberspace. This logical infrastructure provides the foundations for transmitting information across large distributed networks and the creation of online social communities for collaboration and sharing.

Integration refers to the fusing of two or more entities in a coherent and unified manner. There may be combinations across different dimensions of cyberspace. For example, cyber-cyber, social-social, physical-physical, cyber-social, cyber-physical or social-physical. Integration in cyberspace may refer to the coming together of multiple entities in cyberspace under a uniform standard protocol or the fusion of data flowing from multiple interacting sources using a global schema and mapping.

Intelligence refers to the self-enabling ability to acquire process, interpret and apply knowledge. The seamless integration of knowledge across the cyber-social-physical spaces precedes large-scale dynamic collection and processing of information, which allows for self-learning, adaptive behaviours and self-evolution of entities in cyberspace. Intelligence exists across all dimensions of cyberspace in forms of human intelligence (Johnson and Bouchard, 2005), machine intelligence (Legg and Hutter, 2007) and emotional intelligence (Picard, Vyzas and Healey, 2001).

Interconnectivity provides the structure for connecting and linking with other entities in cyberspace while interactions provide communication support between entities that lead to the seamless creation of data and information streams in cyberspace. Cross-domain data are fused by integration which can further be analysed for actionable intelligence.

Over the last twenty years, the majority of the global human population has become actively involved in the production and consumption of information. The interaction, interconnection and integration of entities in our physical and virtual worlds have created a new platform for dynamic information exchange via a distributed communications network. Thus, these interactions, in turn, created logical interdependences between interacting and inter-connected elements.

Data is continually generated in all dimensions of cyberspace through interactions, activities, relationship formation and events. Interactions are achieved through effortless ease of access to cyberspace while interconnections are established in local and global domains during which real-time data is collected, sanitised and stored for future intelligence analysis. Activities in the physical, social, real-world or cyber dimensions are initiated by entities within contextual domains creating historical maps of event-based correlations (Hong *et al.*, 2003). Relationships are formed based on common social, cultural, political, economic or technological beliefs creating associations between entities. Information collection, processing and storage are possible through high-performance digital devices that self-enable knowledge creation and dissemination. The social space is a complex interactive structure made up of individuals, organisations and entities connected by one or more specific types of inter-dependencies such as friendships, interests, likes, dislikes, sexual relationships, political, economic and religious affiliations, relationships of beliefs, kinships, etc. Widely accessible and open source social technologies are increasingly being used by humans to share thoughts, perceptions, opinions and beliefs about real-world events in real-time. Social media platforms like Facebook, Twitter, YouTube produce streams of information capable of creating proactive intelligence. (Debatin *et al.*, 2009);(Lau, Xia and Li, 2012);(Hernández *et al.*, 2016). Cyberspace creates a platform for individuals to operate multiple personalities with which they can share true opinions freely, beliefs

and associations (Personas and Cyber-Personas) (MOD, 2016). A typical example is the formation of online hacker communities', e.g. hacktivists, cyber soldiers as a platform to plan, coordinate and execute cyber-attacks.

Data in the social space exist as a series of interconnected events that captures the real-world personas as well as the cyber personas of online users. Such data are open source and publicly accessible
The physical space consists of real-world dimensions that are mapped into cyberspace. These could be the digital devices through which cyberspace is accessed or real-world perceptions of geographically distributed cyber entities. Data collection is enabled by using sensors, actuators and context-aware networks (Ning *et al.*, 2016). Methods for collection, storage and knowledge extraction on Human-Machine interaction data being generated in real-time for example; network traffic flow (Cleveland and Sun, 1995) can be used to monitor network access (Eterovic *et al.*, 2014), spot malicious behaviours such as denial of service attacks (Jin and Yeung, 2004; Chen and Hwang, 2006), build models for automating security protocols (Sofiyanti, Fitmawati and Roza, 2015) and understand network behaviour (Münz, Li and Carle, 2007; Bou-Harb, Debbabi and Assi, 2015). Typically, software applications hosted on the world wide web could be configured to collect and store interaction data and user data (Tsai and Chan, 2007). End-user behaviour could be inferred using event logs for evidence gathering (Singh and Roy, 2010) and anomaly detection techniques (Breier and Brani\textbackslashvsová, 2015). Activity log data collected from mobile devices or activity devices could be used in health intelligence.
Additionally, data in the real-world dimension such as GPS tracking data, weather data and transport traffic data could be integrated for informed intelligence to trace events such as cyber terrorism or cyber activism (Whiting *et al.*, 2015). Due to the sensitivity of information generated in the physical space and the risk involved with activities around it, data on this domain is not open or publicly accessible. This major challenge could hinder efforts at developing generalised techniques for knowledge extraction (Baggili and Breitinger, 2015).
The decentralisation of networks creates an open market for ideas and beliefs that are often shaped by immediate surrounding events. The complexity of the relationships between real-world events and cyber events provides a new dimension for analysis. Historical and current correlations between social, political, economic and cultural events (Gandhi *et al.*, 2011a) could provide actionable intelligence for cyberspace operations. Additionally, global financial events generate profound effects across various dimensions of cyberspace and the real-world (Pandey and Snekkenes, 2014). As a result, traditional protests or rallies are frequently organised and executed in cyberspace (Olsen, 2013) by grievance-bearing groups. Commodities trade for example the global prices oil and gas, gold cobalt; may also affect global events in the cyberspace and the real-world (Bronk and Tikk-Ringas, 2013).
Similarly, in a contest of competing narratives, the concept of 'news' becomes relative and what is accepted as truth could be considered important when dealing with global events. There is sufficient evidence to prove that useful intelligence can be gathered from a systematic analysis of unstructured open source information especially information from the media (Sousan *et al.*, 2010). For example, studying past social events in news media and publications to explore factors in the social and cultural dimensions that act as antecedents to cyber incidents (Sharma *et al.*, 2010).These enablers of the cyber-physical-social ecosystem are also enablers of the various components of the theory proposed in this thesis.

2.8 CONCEPT OF OPERATIONS

This section defines the CONOPS associated with the proposed approach. Here the researcher lays out a series of sub-framework that characterize the proposed system. This section is a review of how a set of frameworks can be employed and integrated to achieve the parts of the research aim and objectives.

2.8.1 What is a cyber-attack

To reach a definition of 'cyber-attacks' that captures the complexity of the activity, it is essential to the scope of the problem of cyber-attacks. Cyber-attacks are propagated over information communication networks with the intention of manoeuvring or disrupting the natural order of things in the selected network. The aim is to manipulate or alter the medium of communication or the rational logic of the system through malicious means. Once a system is infected with malicious code, the adversary gains control and can remotely take control of the system to perform unwanted tasks within the network. Consequently, if an adversary has access to a single node in the network, the infected node can also be used to affect other nodes in the network.

To manipulate a system, the adversary would expect some flaws or bugs in the applications, software or hardware in the network which he/she hopes to exploit. The motives of the adversary and the means of attack execution provide variety in the definition of a 'cyber-attack'.

Generally, the various definitions of cyber-attacks in research (Gervais, 2012; Hathaway *et al.*, 2012; Klimburg and Mirtl, 2012; Barnett, Smith and Whittington, 2014; DCDC, 2015a), summarize a cyber-attack as any offensive attempt by individuals, groups, organizations or nation states targeted at information and communication networks by any malicious means. This definition highlights some essential characteristics of a cyber-attack.

2.8.2 The Cyber-Attack Kill Chain

Lockheed Martin first published the cyber kill-chain as part of an intelligence-driven cyber defence model (Hutchins, Cloppert and Amin, 2011) for identification, prevention and mitigation of cyber intrusions in target networks. Cyber-attack kill-chains have over the years evolved to a sophisticated model for incident response, digital forensics, malware analysis and understanding the activities in various types of cyber-attacks from the attacker's perspective. Inherently, modelling the kill-chain is modelling and analysing the offensive actions of the adversary. It is important to note however that the cyber kill-chain is a circular and non-linear model where the adversary iterates over many steps, making continuous lateral movements inside the target's network. The cyber kill-chain model states that, for an adversary to carry out a successful attack, he must follow six necessary steps;

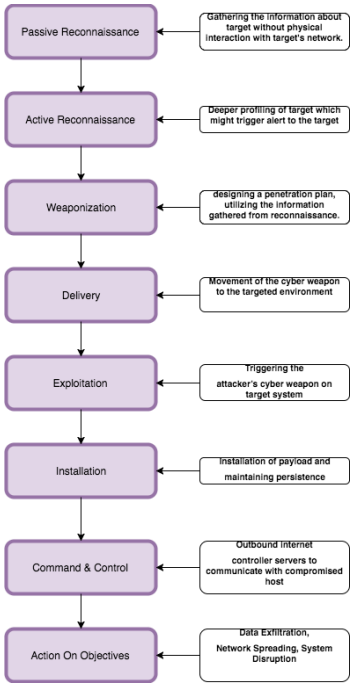

Figure 2-1: The Cyber Kill-Chain Process

The sequential chronological occurrences of events in the kill-chain provides an experimental sequence to event identification in the theoretical development.

2.8.3 Attack Trees

Attack trees are implemented as probabilistic models showing the path which a target or an asset may be attacked. The attack trees introduced by (Schneier, 1999), what sorts of attack to expect, are useful for understanding the goals of an attack, who the attackers are and where to best deploy resources to mitigate attacks. Attack trees represent attacks and counter-measures in a tree structure where the root node is an attack goal and the leaf nodes are possible attack vectors. The nodes represent different ways to achieving the same goal (at the root node) or different steps to achieving the same goal. Attack trees are increasingly implemented in practice due to their ability to quantitively measure the effectiveness of threat analysis and improve decision making around cyber threat monitoring and prevention (Mauw and Oostdijk, 2006; Roy, Kim and Trivedi, 2010).

2.8.4 The Multi-Dimensional Cyberspace

In this section, we discuss the components of a new three-dimensional model for characterising cyberspace. The new characterisation includes the physical dimension and the social dimension as presented by (Clark, 2010; United States: US Army, 2010; Klimburg, 2011). The economic dimension is a new dimension added to capture the effects of political, financial and cultural indicators on cyber incidents as suggested in (Gandhi *et al.*, 2011a). In relation to the cyber-attack kill-chain, the multi-dimensional cyberspace provides transverse movements of events that support channel hopping of the sequential events in the kill-chain.

2.8.4.1 The Physical Dimension

Ning et al. (Ning et al., 2016) describe the physical dimension as a combination of the real world as perceived from a linear dimension and physical objects as they exist in cyberspace. Although Clark and Klimburg (Clark, 2010; Klimburg, 2011) restricted the physical dimension to all physical objects and hardware mapped into cyberspace via sensing and digitisation, some studies have advocated the real-world association to entities in cyberspace. For example, a physical object can be digitised and mapped into cyberspace with a unique IP address. However, it is also possible to map all IP addresses to a physical geo-location in the world which has attributes like weather, traffic, political and economic events, financial markets, commodity markets, economic indicators etc. Therefore, the physical dimension is not limited to include the hardware devices that support access to cyberspace but also attributes that define geographical characteristics of these entities and technical infrastructure that defines the structure of connectivity between them.

- ***The Network***: Sensors and activity monitor support activities on the network layer. Interconnected computing devices in which these devices are perceived and controlled by sensors and actuators to establish a communication link between them. These include all the hardware (servers and computers), infrastructure (wired, wireless, and optical), physical connectors (wires, cables and routers) and digital devices interacting in cyberspace.
- ***The Real-World***: The geographic or real-world components of cyberspace refer to the physical location of elements of the network. These are the characteristics that define a physical location of a network object. It is essential to differentiate the real-world components from activities on the economic dimension. While activities in the economic dimension relate to events that affect economic, political, social and cultural dynamics, real-world components are seen as a naturally occurring phenomenon. Therefore, the geographical components are not restricted to the physical dimension.

2.8.4.2 The Social Dimension

The social dimension comprises of the persona and cyber-persona components (Clark, 2010; Klimburg, 2011; United States Defense Force, 2013). The social dimension is an integration of social attributes and inter-relationships owned by human beings and another related physical, cyber entities in cyberspace. The personas and cyber-personas characterise human activities, social events, behavioral conventions, political administrations, public services for human social participants in cyberspace. Social interconnectedness establishes the existence of the social dimension by enabling the creation of online communities characterised by the active social presence, social participation, relationship creation, and collaboration and belief affiliations.

- ***The Cyber-Persona***: The components of the cyber-persona define online identities and characteristics of the people in a network as they exist and interact in cyberspace. These elements refer to the digital footprints of individuals as they interact with other aspects of cyberspace. Individual cyber-personas are attributed directly to an actual person or persons incorporating some cyber characteristics, e.g., email address, IP address, social accounts, etc. A single individual may possess multiple cyber personas, and a single cyber-persona can have multiple users.
- ***The Persona***: The persona refers to the offline mode of the cyber-social dimension (Ning et al., 2016). The persona components define the identities and characteristics of the people interacting in a network from a real-world perspective and the characteristics that define them

as they exist in the real world. For example, a person may have attributes like a geographic location, a financial service provider, a home address or a work address.

2.8.4.3 The Economic Dimension

The economic dimension characterises cyberspace as an enabler of our socio-political-economic-cultural existence. The socio-political-economic-cultural dimension is characterised by activities attributed to human beliefs, norms, laws and co-existence. As a result, it captures the state of human experience as they exist as individuals or communities. It reflects how activities on the real-world economic and political hemisphere affect the events in cyberspace (Sharma *et al.*, 2013). Current research has made attempts to quantify the association between socio-political-economic-cultural factors and events in cyberspace. For example (Cavusoglu, Mishra and Raghunathan, 2004) studies the effects of various types of cyber-attack announcements on stock market prices and (Gandhi et al., 2011; Karatzogianni, 2008) identifies the effects of social, economic, cultural and political indicators on the occurrence of cyber-attacks. Although there has been little focus on the direct impact of stock and commodity market price fluctuations on the event of cyber-attacks, there is sufficient evidence supporting their influence across multiple layers in cyberspace (Sharma *et al.*, 2010; Bollen and Mao, 2011; Sanzgiri, Hughes and Upadhyaya, 2013; Hernández *et al.*, 2016). Recent activities in cyberspace also prove how economical (Olsen, 2013), political (Leigh and Harding, 2011; Harding, 2014; Zetter, 2014) and cultural (Bartlett, 2015) dimensions of society lead to cyber-incidents.

- ***The Broadcast***: The realistic real-time content generation of real-world incidents supports the existence of this layer. This layer also relates to information about events in the real world. A unique characteristic of this layer is its real-timeliness and newness of information generated. These are facts or 'knowns' and include information from certified media sources. While some may argue that Twitter, Facebook, and other social media platforms can also act as news dissemination actors (Kwak *et al.*, 2010), the news is seen to be biased in such aspects by the uncertainty of the information sources.
- ***The Markets***: This layer captures the effects of money, finance and trade on our modern existence and how they may act as antecedents to cyber defence events. Therefore, we seek to identify what market forces influence cyber defence incidents. Markets in this context refer to all financial, trade and money context of the real world reflected in cyberspace. For example, fluctuations in exchange rates, commodity prices, stock prices, crude-oil prices, money supply, money demand, inflation rates, deflation rates may have direct or indirect impacts on economically or politically motivated cyber-attacks. Moreover, (Bollen and Mao, 2011) proves that elements in cyberspace can influence stock market prices as well as market valuations being influenced by denial-of-service attack announcements on the market valuation of firms (Cavusoglu, Mishra and Raghunathan, 2004). Similarly, (Bronk and Tikk-Ringas, 2013) explains how events in the oil and gas sector affect the propagation of cyber incidents.

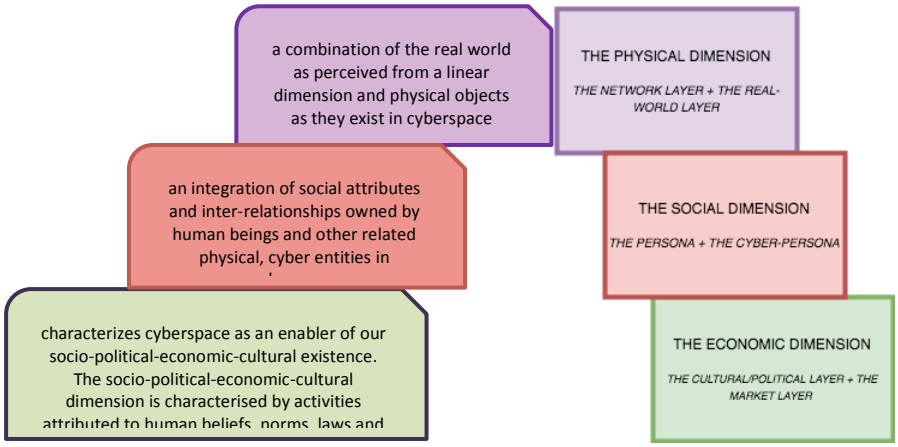

Figure 2-2: The Multi-Dimensional Cyberspace

In conclusion, we can refer to cyberspace as a multi-dimensional space with multiple independent dimensions that make up the whole. In an attempt to accurately represent cyberspace, the researcher denotes cyberspace as Ω. Where Ω refers to all data and information flowing across the various dimensions of cyberspace. Therefore cyberspace:

$$\textbf{Cyberspace} = \Omega = \begin{pmatrix} \boldsymbol{Physical\ Dimension} \\ \boldsymbol{Social\ Dimension} \\ \boldsymbol{Economic\ Dimension} \end{pmatrix}$$

Equation 2-1: Multi-Dimensional Representation of Cyberspace: Compact View (Source - Author)

Given a multi-dimensional cyberspace with J dimensions, let $\Omega_1 \ldots . \Omega_J$ represent indicators from the corresponding dimensions of cyberspace. Therefore, we can represent data from these dimensions as:

$$\boldsymbol{Cyberspace} = \begin{bmatrix} \Omega_1 \\ . \\ . \\ . \\ \Omega_J \end{bmatrix}_t$$

Data from each dimension can further be decomposed into an (n X m) matrix where m is the number of features or indicators observed within each dimension and n is the number of chronological observations for each indicator from time $\boldsymbol{t_1} \ldots \boldsymbol{t_n}$. Therefore, Ω_1 is an (n X m) matrix where m is the number of features for the first identified dimension of cyberspace and n is the number of chronological observations recorded at the given time interval.

2.8.5 Data and Information in Cyberspace

This section provides a conceptual breakdown of data and information available as evidence sources in cyberspace.

2.8.5.1 Data From the Physical Space

The physical space acts as a gateway to all cyber-attacks. Data analysis on the physical space, most notably the network layer seems to have dominated literature in the past decade. Various types of data that are extracted from the network layer of cyberspace have extensive coverage in literature.
Time series models are commonly applied to network traffic data to predict network flow (Barford *et al.*, 2002; Kim and Reddy, 2008), to spot denial of service attacks (Kim & Reddy, 2008), in building intrusion detection systems (García-Teodoro, Díaz-Verdejo, Maciá-Fernández, & Vázquez, 2009) and classifying network traffic packets (Mahoney, 2003). Moreover, weblog data is another source of data for proactive cyber defence models to network traffic flow data. For example, probabilistic latent semantic analysis (Tsai & Chan, 2007) detects cyber defence threats in weblog data files. End-user behaviour can also be inferred using event logs for evidence gathering (Singh & Roy, 2010) and anomaly detection techniques (Patcha & Park, 2007). Similarly, the use of classification and correlation (Gabra, 2014) techniques on IDS alert logs support cyber forensics and real-time cyber threat assessments. Consequently, data and information from sensors such as the Internet of things (Atzori, Iera, Morabito, & Nitti, 2012) also fall within the physical layer of cyberspace. Data on the physical layer may also include some real-world attributes, e.g., weather data, GPS tracking data and some traffic data to trace events such as cyber terrorism or cyber activism (Whiting et al., 2015).
Unlike the social layer, one major disadvantage of using data on the physical layer is that legalities and access guard most of it is not as open. Consequently, attempts to perform analysis on this layer has always been restricted to contexts analysis. Therefore, results are not readily generalizable (Baggili and Breitinger, 2015).

2.8.5.2 Data From the Social Space

The social space is a complex interactive structure made up of individuals, organisations and entities connected by one or more specific types of inter-dependencies such as friendships, interests, likes, dislikes, sexual relationships, political, economic and religious affiliations, relationships of beliefs, kinships, etc. Widely accessible and open source social technologies are increasingly being used by humans to publish thoughts, perceptions, opinions and beliefs about real-world events in real-time. Social media platforms like Facebook and Twitter are a rich source of data for proactively countering cyber-attacks. For example, to encourage online user privacy awareness, (Debatin *et al.*, 2009) uses Facebook data to measure users' online exposure to vulnerabilities. By studying users' ritualisation, routines, gossips, rumours, this study could examine the relationship of Facebook privacy issues with invasion of privacy.
Additionally, (Bou-Harb, Debbabi and Assi, 2015) crawled cyber-crime related messages from Twitter and context-related blog articles to identify notorious hacker groups and their related twitter feeds, blogs and blog activities. Also, cyber terrorist groups have also been spotted using text mining and analysis techniques. Hernández (Hernández *et al.*, 2016) employs user sentiment analysis on a daily collection of tweets in two different contexts; Twitter as a platform for expression of user views and Twitter as a platform to present content related defence threats on the web. Statistical analysis of these data is used to predict the possibility of a future attack.
Data on the social layer of cyberspace exist as a series of interconnected events that captures the real-world personas as well as the cyber personas of online users. Consequently, data and information from social media platforms such as Facebook, Twitter and micro-blogging platforms that allow users to freely share views, perceptions, and opinions about related matters, feeds into cyber-attack counter models. One distinct advantage of using data from the social layer is the free and open accessibility it provides.

2.8.5.3 Data From the Economic Space

The complexity and the interconnectedness of real-world events and cyber events uniquely characterises the economic space. Similarly, time series models are also equally applied to stock market data as generated in real-time. Although there is little research on the effects of stock market prices on cyber defence events, research has covered the integration of stock market data with data from other layers of cyberspace (Bollen and Mao, 2011).

Sufficient evidence supports the extraction of cyber threat intelligence from a systematic analysis of unstructured open source information, e.g., text/news articles (Tsai & Chan, 2007). Similarly, studying past social events in news media and publications reveal factors in the social and cultural dimensions that act as antecedents to cyber-attacks (Whiting et al., 2015). The usefulness of financial prediction markets in information defence risk management is another example of this concept (Pandey & Snekkenes, 2014b) with methods such as (Pandey & Snekkenes, 2014a) suggests.

2.8.6 The Socio-Physical-Economic Cyber Kill Chain

The cyber-kill chain model presented by (Hutchins, 2011) presents a phased approach to orchestrating cyber-attacks. The cyber-kill chain model identifies the steps which the adversary must complete for a successful attack on victim's networks and is part of the intelligence-driven defence model for identification and prevention of cyber-attacks. The chain begins with intelligence gathering and ends with the cyber-attack. The cyber kill-chain frameworks prove that cyber-attack happens in phases and stopping the adversary at any of the phases breaks the attack chain therefore significantly reducing the probability of success.

In a multi-dimensional cyberspace, as discussed in section 2.8.42.8.4, the deconstruction of the phases of the cyber-kill chain across the dimensions of cyberspace produces a socio-physical-economic kill-chain where each activity on each phase of the kill chain is uniquely identified as occurring on one or more dimensions of the multi-dimensional cyberspace. The socio-physical-economic kill chain framework integrates two logical sub-frameworks; the cyber-attack kill chain (Hutchins, 2011) and the multi-dimensional cyberspace as discussed above.

The cyber-attack kill chain introduced by Lockheed Martin represents a generalised sequential process that cyber-attacks follow from conception to execution. The resulting combined framework outlines the various life-cycle phases in a cyber-attack from cyber-attack trigger to cyber-attack execution as they happen across the dimensions of cyberspace.

Predictive cyber defence models should incorporate the underlying factors or antecedents that lead up to the occurrence of a cyber-attack. Therefore, a multi-dimensional structural approach capable of representing the complex dynamics of cyberspace is used.

The socio-physical-economic cyber-kill chain presents the cyber-attack kill chain from a multi-dimensional perspective aimed at understanding the cyber-attack lifecycle across the various layers of cyberspace, from the adversary's perspective. This framework aims to spot the indicators of cyber incidents at every stage of the kill chain. One way to determine the indicators of each event on the kill-chain is to plan a cyber-attack from the attacker's perspective by analysing the tasks an attacker must carry out to complete each phase. In this research, we refer to phase indicators as evidence in the data characterising the kill-chain phase. Consequently, these evidences are unique for different types of cyber-attacks.

2.8.6.1 The Antecedents

There are events and situations in the environment that precedes the behaviour of a cyber-attack. A behavioural epistemology considers these preceding events and situations as antecedents. Antecedents in cyberspace prompt the occurrence of cyber-incidents in that they act as precursors. Consequently, we seek to identify what conditions in the dimensions of cyberspace serve as antecedents to the behaviour of cyber-attacks. Additionally, this approach to applied behavioural sciences leads to the identification of motivations of cyber-incidents as it reveals what antecedents' adversaries respond to.
Researchers explain that cyber-incidents are associated with Social Political Cultural and Economic events-which are the most promising antecedents of cyber-incidents in cyberspace (Gandhi *et al.*, 2011a; Ning *et al.*, 2016). SPEC analysis of cyber events provides a medium for analysing the motivations behind certain cyber-attacks. Socially motivated cyber-incidents are seen as resulting rivalries between individuals or groups over incompatible goals, scarce resources or control. Culturally motivated cyber incidents result from differences in ideologies, beliefs and a sense of a territorial way of life that is unique to certain groups. Economically motivated cyber-incidents are preceded by economic, corporate or financial situations that lead to uprisings by certain groups. Overcoming the challenge of providing empirical evidence of the association between cyber-incidents and SPEC factors is an on-going research milestone for cybersecurity.
Consequently, there is currently little research attributing the antecedents of cyber-incidents to any dimension of cyberspace; however theoretical research on the subject (Gandhi *et al.*, 2011a) have claimed these antecedents to be clustered on the physical dimension-the real-world layer and the economic dimension-the media and the market layers. One attempt to quantitatively prove the association between SPEC events and cyber-attacks is presented by (Gandhi *et al.*, 2011a) use methods such as country level correlation and quadratic assignment procedure to analyse the relevance of social, cultural and economic factors to the occurrence of a cyber-attack.

2.8.6.2 The Reconnaissance Phase (RP)

The reconnaissance phase is characterised mostly by activities on the social and network layers of cyberspace. The reconnaissance phase begins with a target or target systems about which an adversary must gain the maximum amount of information possible by any means necessary. This involves the identification, selection and profiling of a target. The 'target' may be an individual/individual or an organisation and may be passive or active. Passive reconnaissance is carried out un-noticed by the target. However, active reconnaissance may alert the target that something may be going on. The various means by which information about targeted systems may be acquired includes:

- ***Open source reconnaissance:*** The adversary seeks to gather intelligence from publicly available information collected, exploited and used with a time-dependent advantage. For an adversary, this information may include public information on target's employees, company profile, network domains and IP addresses, people associated with the target, social media presence, news articles and broadcasts, government documents freely available online. As most of this information is usually given to the public via the world-wide-web, reconnaissance efforts here are done on the social layer of cyberspace. The adversary surfs the web for publicly available information on target systems. For example, an adversary may infer that a target company runs a Windows-based company-wide system having discovered over twenty job ads for 'A Windows Administrator' or 'C++ or C# Developers' over the past three years. An adversary may also infer the level of cybersecurity importance for a given company by observing the percentage of its staff on LinkedIn who have some sort of security training or awareness.

- ***Network Reconnaissance***: The adversary also seeks to create a near approximate blueprint of the target network. For an adversary, these include the number of active IP addresses on target's network range, the allocated IP addresses that are currently active and the operating systems and services running on each host in the target's network range. Given that in most cases, the adversary can only obtain this information through physical access to the target's network, most of the activities can be detected on the target's network layer. For example, an adversary may perform a ping sweep or a port scan on a target network which can be spotted with techniques outlined in (Adams and Heard, 2014). However, adversaries may also take proactive steps to stay hidden, disguised or manipulate system logs to erase any trace of their digital existence on the target network.

The passive social reconnaissance and the active reconnaissance are usually apart by a given time lag (t-p) with the social reconnaissance as the preceding event. Here, we categorically define a complete 'cyber-attack reconnaissance' as a social reconnaissance event followed by an active network reconnaissance at a time interval k. Given the possible data characteristics of the social and physical layer with the known activities of the adversary at the stage, it is possible to characterise what patterns represent the event of the reconnaissance stage. Although the social reconnaissance may not be a strong indicator of the start of the attack kill chain, a social reconnaissance in combination with a network reconnaissance or a network reconnaissance by itself, typically flags off the start of the cyber-attack kill chain. The reconnaissance phase of the cyber-attack kill chain is a targeted research project with the aim of identifying, selecting and confirming operational and security loopholes in a cyber network. Combined information on target systems exposes vulnerabilities of the target and provides intelligence on how best to carry out a certain attack mission with the target's system framework. Based on a certain set of target network system and attack resources, an adversary can infer a probability of success for a given number of attack vectors and attack paths.

2.8.6.2.1 Identifying Indicators of The Reconnaissance Phase on The Cyber Kill-Chain

The ability to identify characteristic evidence of an active reconnaissance requires a prerequisite knowledge of the activities an attacker must undertake to ensure a successful reconnaissance. As earlier mentioned, reconnaissance is simply information gathering that supports the attacker's attempt in future stages of the kill-chain. Active reconnaissance techniques such as ping sweeps, banner grabbing, port scans, vulnerability scans have unique indicators on the target's network. In most cases, active reconnaissance requires direct interaction with the target's network.

Port scans and ping sweeps represent a sizable portion of active reconnaissance techniques employed by attackers. Pingsweeping is a technique for mapping a range of IP addresses to live hosts while port scanning is a technique for discovering weaknesses and vulnerabilities on live hosts by sending port probes. Although network administrators sometimes use port scans for exploration or testing, ping sweeps and port scans often refer to scans carried out by malicious attackers to determine weaknesses in a target's network.

For this research, we utilise Bailey, Roedel and Silenok's (Bailey Lee, Roedel and Silenok, 2003) three basic classes of scans based on the pattern of attacker movement within the network, target's destination and target ports. A vertical scan targets several destination ports on a single host. A vertical scan is a function of the number of unique destination ports and the number of unique source IP addresses communicating in a network per unit time. Consider the following equation,

$$\frac{Number\ of\ distinct\ Source\ IP\ addresses}{Number\ of\ distinct\ destination\ ports}$$

Equation 2-2: Vertical Scans in Network Traffic (Bailey Lee, Roedel and Silenok, 2003)

For normal traffic, for any given time frame, we expect this figure to be as close to the mean or 1 as possible. On the other hand, consider a horizontal scan, which targets the same port on several hosts. The horizontal scan is, therefore, a function of unique destination IPs within the network and unique destination ports connected to within the network.

$$\frac{Number\ of\ distinct\ Destination\ IP\ addresses - Number\ of\ distinct\ Source\ IP\ addresses}{Number\ of\ distinct\ destination\ ports}$$

Equation 2-3: Horizontal Scans in Network Traffic (Bailey Lee, Roedel and Silenok, 2003)

Finally, block scans combine both vertical and horizontal scanning techniques in large sweeps of IP address range. Moreover, (Nam and Kim, 2006) explicitly describes the scanner detection rule as:

$$\frac{Number\ of\ distinct\ Source\ IP\ addresses\ attempting\ to\ connect\ to\ a\ source}{Number\ of\ conenction\ attempts\ to\ distinct\ IP\ addresses\ by\ a\ single\ source}$$

Equation 2-4: Block Scans in Network Traffic (Bailey Lee, Roedel and Silenok, 2003)

In this regard, we can define a port scan as a function of the number of unique sources IP addresses and the number of unique destinations addresses. Finally, (Kinable, 2008) presents a hypothesis that measures the connection fail ratio of a network at any given point in time as a function of the number of inbound connections and outbound connections. The CFR of a network is given as:

$$\frac{Number\ of\ outbound\ connections - Number\ of\ inbound\ connections}{Number\ of\ outbound\ connections}$$

Equation 2-5: Connection Fail Ratio of Network Connections (Bailey Lee, Roedel and Silenok, 2003)

Attackers, however, are well aware of these detection techniques and employ some methods to avoid detection. This presents a challenge for establishing reliable results based on the techniques used above. One common way to avoid detection is to increase the amount of time between probes since most detection systems perform analysis within a time-spaced window. This way, the attacker can spread out the attack thereby normalising the overall effect on the network over a long period. An attacker can also perform coordinated attacks using IP decoys appearing as different scans originating from several IPs. Identifying decoys in a network analysis the number of source IPs targeting the same set of hosts where these source IPs are within the same network (Bailey Lee, Roedel and Silenok, 2003).

2.8.6.3 *THE WEAPONISATION PHASE (WP)*

The weaponisation stage is characterised mostly by activities on the social and network layers. The goal of this phase is to develop a cyber weapon that executes the intended cyber-attack on a target. The payload developed in this stage is based on the outcomes of the reconnaissance stage. A cyber weapon is a collection of one or more payloads, each of which exploits a specific vulnerability of a target on the network layer, with the goal of disrupting, destroying or manipulating its target. The forensics of the attack weapon is comprised of various characteristics; which depends on the adversary's intended purpose, to flood communication, gain access or gain control.

The weaponisation can also be called 'the blind spot', as it happens outside the reach of analytical purview. In addressing this limitation, this model uses an inference-based approach, based on prior network knowns. Information from previous patterns in network behaviour in combination with common knowns feed the model at this stage. The goals are:

a. To understand the pattern of vulnerability and attacks within the target's network.
b. Conduct a zero-day and a non-zero-day vulnerability assessment on each node in the network and the target network.
c. Construct a weaponisation probability tree for each node in the network and the whole network.

The weaponisation phase probability tree is a combination of tree analysis and inference-cased techniques used to estimate probability functions. For each node in the target network, a weaponisation phase probability tree or attack weapon probability tree is constructed as seen in figure 2-3 below.

a. Construct a node feature table. This includes all hardware, software, application, usage, access features, open ports and network features of a machine linked in a network.
b. Calculate the probability of exposure of each node. This involves extracting the keywords from each node feature and passing them through vulnerabilities databases, exploit databases and cyber-attack databases for a match.
c. Weight each return against the total number of unique versions of applications, software running within the network.

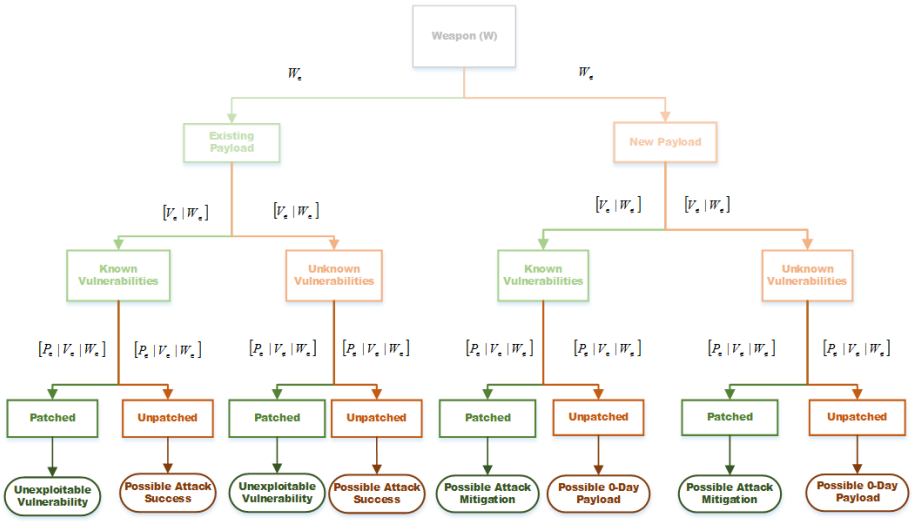

Figure 2-3: Weaponisation Probability Tree (Source: Author)

A recommended starting point for the weaponisation inference is to gain an understanding of what an adversary needs to access a target network based on the adversary's intention. The adversarial tactics, techniques and common knowledge framework describe multiple approaches to designing attack

vectors with which an adversary may take to fulfil intentions. Based on the number of tools available to the adversary and the potential ease of access to these tools, an estimated weight of attack technique percentage probability is created. For Example, the table below shows adversary techniques and corresponding weighted percentage probability of use.

S/N	Attack Plan decision scenario	Number of Available Attack Vector Techniques	Weighted Percentage Probability
1	Advanced Persistent Attack	25	15%
2	Escalate Privilege	14	8.7%
3	Evade Network Defences	32	19.8%
4	Access Database or Credentials	8	5%
5	Discovery	16	10%
6	Lateral Movement	14	8.7%
7	Execution	16	10%
8	Collection	11	6.8%
9	Data Exfiltration	9	5.6%
10	C&C	16	10%

Table 2-1: Attack Plan Vectors. Source (Kick, 2014)

Another approach to weaponisation inference is based on scan statistics. Historical evidence shows that understanding network scan statistics is a critical aspect of any cyber situational awareness model (Wu and Shao, 2005; Kim and Reddy, 2008; Adams and Heard, 2014). Network scan analysis is especially useful for detecting malicious activities within a network. Although (Panjwani *et al.*, 2005) demonstrates that over 50% of cyber-attacks at the time are not necessarily preceded by some form of scanning activity, during a network reconnaissance stage, the combination of a port scan and a vulnerability scan or a port scan, a vulnerability scan and an ICMP scan is a strong indicator of possible attacks (Panjwani *et al.*, 2005). Based on 'what was scanned', defenders may produce a chronological sequence of scan probes for dynamic vulnerability assessment. For each observed vulnerability and port scan, an estimate of the probability that the vulnerability would be exploited is created.

In addition to the class of scan performed (port scan, vulnerability scan, ICMP scan), another feature that may prove useful in determining the adversary's weaponisation tactics is the type of scan performed (Bhuyan, Bhattacharyya and Kalita, 2011). Vertical scans perform multiple scans on a single host. This may occur when an adversary is interested in a specific machine in a network. Horizontal scans perform a single scan on multiple hosts. In this case, the adversary aware of an existing vulnerability and tries to detect susceptible hosts in a network. Block scans perform both vertical and horizontal scans in a sweep of the entire network. In this case, the adversary creates a pre vulnerability assessment of the target's network. The combinations of the class and type of network scan performed create some scenarios for adversary intent.

The use of scan statistics for weaponisation inference is limited in that only adversary activities in network reconnaissance can be spotted. An adversary may utilise other forms of reconnaissance, for example, social engineering, physical access, insider threats. Also, an adversary may not always need a network reconnaissance to execute an attack. Once these the occurrence of the combination of these events have been established, defenders may assume that the adversary may have found a vulnerability which will be exploited at a future time 't'. The usefulness of inferring the adversary's

weaponisation end-game is subject to time, particularly, the timing of the cyber-attack (Axelrod and Iliev, 2014) and the nature of the vulnerability lifecycle (Marconato, Nicomette and Ka??niche, 2012).

2.8.6.3.1 The Vulnerability Lifecycle

These categories are based on the strategic timing of activities carried out by the adversary to succeed and by the defender to mitigate.

- Vulnerability Birth Date: A human programming mistake or a loop-hole in application and software source code leads to a vulnerability to be exploited by adversaries.
- Exploit Availability: Adversaries in the deep web society find this vulnerability and create exploits and conduct quick short-term attacks against selected targets.
- Vulnerability Discovery date. This refers to the date the vulnerability is discovered. The vulnerability is discovered either by security experts or adversary hackers. If the vulnerability is discovered by security experts, the incident is reported to necessary security vendors, and details of the vulnerability are released to the public only after a patch has been issued. However, if the vulnerability is first discovered by an individual or a black hat hacker, the vulnerability is most likely to be exploited before its existence is disclosed.
- Vulnerability Disclosure date. This is the date the vulnerability is disclosed to some groups of the public. At this stage, the vulnerability existence is not yet publicized to the public. The exact dates of vulnerability disclosure dates are captured in the Open Source Vulnerability Database.
- Vulnerability Publicity date. Media outlets like CNN, NYTimes, BBC, C|NET, ComputerWorld, HackReads, Hackmeggedon, and eWeek. However, security vendors are already aware of the vulnerability and are likely to release a fix. This, in turn, reduces the motivation of the adversary to attack.
- Vulnerability Exploitability date. This is when an exploit code for the vulnerability is created and publicised on websites. The details of vulnerabilities will be given out on specific websites such as Packet storm or mailing list websites such as NEOHAPSIS.
- Vulnerability Patch date. This is when a patch or fix is available. This is obtained directly from the vendors' websites.

2.8.6.3.2 The Timing Of Cyber Attack

As pointed out by (Axelrod and Iliev, 2014) when an adversary creates a weapon to exploit a vulnerability on a target network, the adversary decides whether to use that payload immediately or wait for a specific time in the future. This wait may be due to some strategy being implemented by the adversary or the weight of stakes involved. The adversary may be individuals, hacker groups or nation states. Per (Axelrod and Iliev, 2014), there is a trade-off for an adversary between waiting until the stakes are high enough to rip the maximum benefits from the use of an exploit and waiting too long enough that the target identifies and patches said vulnerabilities. The adversary's decision on the optimum time to use an exploit against a target is based on certain factors.

a. The adversary's perceived stakes: This is what an adversary or defender stands to lose or gain depending on the success or failure of a cyber weapon. An adversary is concerned with the possible gains or losses if the cyber weapon is used at a given time to exploit a vulnerability while the defender is concerned with the potential gains and losses if the cyber weapon successfully infiltrates and attacks the network.
b. Adversary stealth

c. Adversary Persistence
d. The perceived value of Cyber Weapon
e. The adversary's skill level
f. Defender's Pro-activeness
g. Defender's mean time to recovery.

For the proactive defender, assuming all open vulnerabilities have been patched in time on the way to estimate this time is to evaluate the meantime to exploitation (Leversage and Byres, 2008) or predict the exploitability of vulnerabilities in cases of a zero-day attack (Kao *et al.*, 2015).

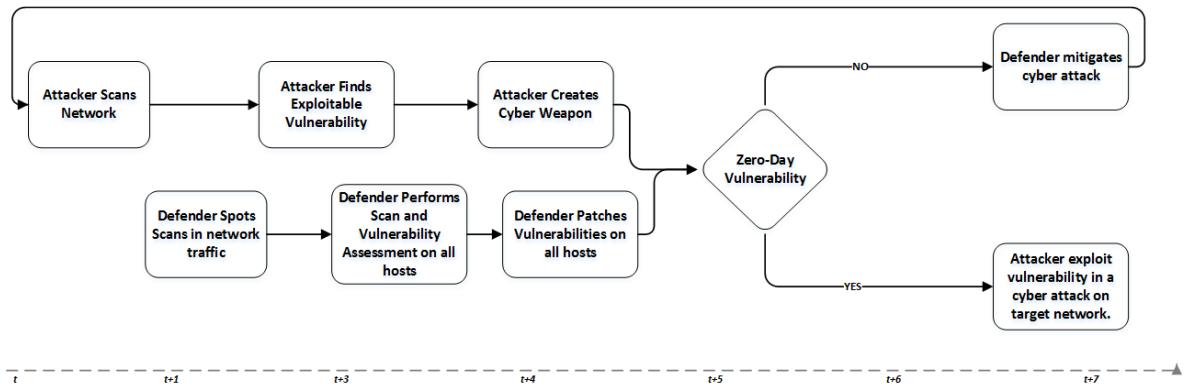

Figure 2-4: Proactive Targeted Defender

This cycle of vulnerability discovery to network attack is slightly different for targeted attacks and random attacks. For targeted attacks as seen in Figure 2-4 above, the attacker is interested in a specific target and therefore scans that target network for any possible open vulnerability. It is only after this scan has been completed and the attacker is armed with critical information on the target network's topology, that a cyber weapon to exploit is created. In the case of targeted cyber-attacks, one major challenge of the weaponisation phase is the minimal knowledge on adversary's strategy. Adversaries' game plans are usually identified and spotted after-the-fact, and defenders are stripped of the luxury of anticipation. However, by constructing a predefined set of 'givens' based on past events in the kill-chain, target profile, vulnerability assessment, an attack weapon probability tree is constructed.

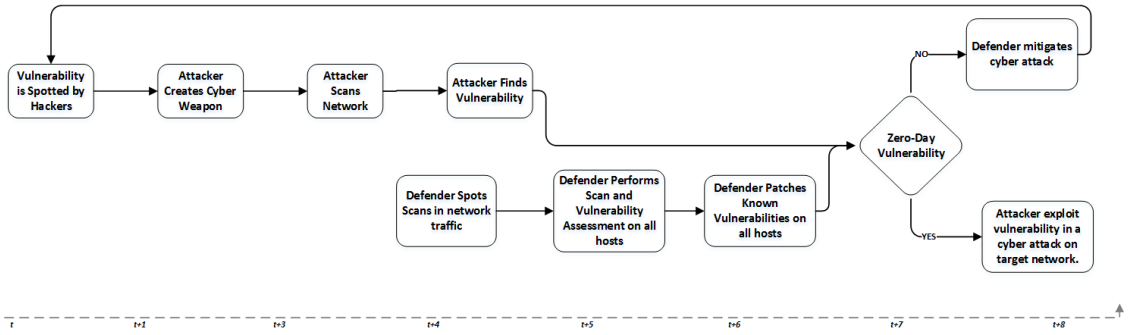

Figure 2-5: Proactive Random Defender

For random attacks with random targets as seen in *Figure 2-5* above, the attacker possesses an exploit and goes searching for any possible vulnerable target. Here, the attacker is well armed with a cyber weapon before discovering the vulnerability on the target's network. The implication of this is a significant reduction in the time available to defenders to spot scan activities and patch vulnerabilities. In the case of a proactive random defender, time t_4 to t_6 is very crucial for both attackers and

defenders. Defenders try to patch vulnerabilities before the attacker has time to deliver the cyber weapon while attackers are trying to go unnoticed in conducting their malicious activities.
However, there is also the possibility of a zero-day attack in which case; the defender is unaware of the existence of said vulnerability. This is difficult to spot as data on vulnerability is only available after the payload has been used, in which case it is no longer considered a zero-day and the adversary's motivation to use it is significantly reduced. Cyber network vulnerabilities are grouped into one of five categories during their existence (Jumratjaroenvanit and Teng-Amnuay, 2008).

- Discovered Zero-day vulnerabilities: here, the adversary discovers an unknown vulnerability in the target's network and secretly works on a payload to exploit that vulnerability while ensuring the existence of the vulnerability is not disclosed.
- Active zero-day Vulnerability: Here, a patch for a known zero-day vulnerability has been released by security or software vendors. However, internal network and systems administrators fail to apply the patch to vulnerable systems a considerable time before the adversary has the chance to utilise the payload.
- Passive zero-day Vulnerability: This is very like the pseudo-zero-day vulnerability except, this vulnerability has not been exploited yet. The system or network administrators have failed to apply the appropriate patch, and the vulnerability has a high risk of being exploited.
- Active Open Vulnerability: Here, system, software and application vulnerabilities have been publicised and the exploits are openly known. However, security and software vendors are unable to produce a suitable patch for said vulnerability in the time leading up to the eventual exploitation of the target's network.
- Passive Open Vulnerability: Here, the vulnerability may be known or unknown to network administrators and security vendors. However, the exploit codes to exploit said vulnerability has not yet been produced.

Zero-day vulnerability is a vulnerability of which a security patch has not been released and is therefore open to exploitation by unknown adversaries. In such cases, the defender may or may not be aware of the existence of such vulnerability, but most likely has little or no idea of how an adversary would design a payload to exploit said vulnerability. When an adversary successfully exploits a zero-day vulnerability before a security patch is released, the target is said to be a victim of a zero-day attack.

2.8.6.3.3 Identifying indicators of weaponisation on the cyber kill-chain

At this point, the attacker has enough information on what is needed to infiltrate the target network. The attacker has found an entry point and goes about developing a cyber weapon to achieve this purpose. Unfortunately, the weaponisation phase is a blindspot in the kill-chain and represents an area where the target has no control of mitigating. However, from a target perspective is possible to probabilistically infer attack vectors from indicators identified in the reconnaissance phase. This method assumes that the target is aware of the activities carried out by the attacker in phase 1. For example, what ports were scanned, which employee's information was searched etc. Based on this information, target constructs multiple hypotheses on feasible attack vector scenarios which in turn splits the kill-chain into multiple plausible paths.
Estimating the likelihood of expectation for different cyber threat vectors is a form risk analysis on a target's network. All plausible attack vectors for a given cyber-attack must be identified and a probability of expectation value assigned to each one. For example, assuming it is known that ports A,

B and C were scanned, and port B is associated with SQL database servers, the target maps all ways an attacker may infiltrate the servers through an SQL vulnerability. This strategy requires prior knowledge of exploits, vulnerabilities and attack vectors associated with the identified servers. A target can also infer the possibility of a zero-day attack as seen in figure 1 above. As a cyber-attack progresses, this strategy reduces the outcome space by assigning a certain amount of probability to each attack vector scenario which in turn provides multiple hypothesis for the attack kill-chain.

2.8.6.4 The Delivery Phase

The delivery phase is characterised mostly by activities on the network and real-world layers. The delivery state involves all activities that support the movement of an exploit from the adversary to the target, or the transmission of the cyber weapon to the target's territory. The delivery stage is very critical to both the adversary and the target as it is the basis of a clean, efficient cyber-attack. The method of delivery can vary from email attachments, USBs to malware injections. At this stage, the defender may estimate the chances that a cyber weapon has successfully infiltrated the system or the defensive capacity of the network by estimating the chances that a cyber weapon will successfully infiltrate the network. The delivery stage is linked to the physical and information layers of cyberspace. On the physical layer, it may be transmitted as malware in scripts that silently insert themselves into a host computer, from which it spreads to the entire network, DNS cache poisoning or as a flood of data packets sent specifically to crash target's network e.g. in Denial of service attacks. On the information layer, delivery mechanisms could be phishing emails or email attachments. Unfortunately, the delivery stage leaves traces of adversary activity on the target network. Therefore, the adversary performs this task anonymously. It is therefore almost impossible to specifically identify who the adversary might be. However, there are constants in a target network that may help defenders predict the occurrence of such delivery mechanisms. Multiple classifier systems use Markov chains in detecting injection attacks (Roy, Singh and Sairam, 2011; Alnabulsi, Islam and Mamun, 2014) and attack propagation methods in networks (Bar *et al.*, 2016). Another approach is a probability-based model to estimate the delivery mechanism based on target network characteristics. It is possible to estimate, given past occurrences of cyber-attacks and target network characteristics what delivery mechanism a network is most susceptible to give the operating system and application platforms running on a network.

The signatures that characterise the delivery stage are based on the permits of information flow within a target network. For example, an insider attack may be able to use a USB stick to deliver malware because the administrative protocols allow the transfer of data from unidentified external storage devices. The task of delivering a cyber weapon to target network may be via wired or wireless communications. A cyber weapon may successfully be delivered via emails because the network has no email or spam filtering available. Sometimes files run without administrative approval which may lead to download and installation of exploit by an unknowing network participant. Once indicators have been observed, to predict the probability of delivery success, all points of entry are first identified within a target network. The characteristics of the delivery stages are assumed to be on the adversary's network layer and the target's network layer.

2.8.6.5 The Exploitation Phase

The exploitation phase is characterised mostly by activities on the network, real-world and information layers. On successful delivery of the cyber weapon, the target completes the required steps that execute the weapon in the target network, thus triggering the exploit. This phase is

characterised by the modification and degradation of the cyber layer in such a way that confidentiality, integrity and availability have been compromised. The objective at this stage is to install and execute the cyber payload. Therefore, certain pre-conditions must be met. Firstly, the target should have running the operating system or software, which the cyber exploit was made for. Secondly, the vulnerability to be exploited must have a status of 'open' at the time of exploitation, i.e. the operating system or system application should not be upgraded in such a way that it renders the cyber exploit useless. Lastly, the delivery must be done in such a way that its movement across the target's network is not spotted by anti-virus or other protection software. In such a case that the cyber exploit is spotted, vulnerabilities are immediately patched, and exploit is rendered useless. At this stage, an adversary uses an exploit kit to mitigate the chances of failure of one specific exploit on a vulnerability. An exploit kit is a combination of multiple exploits crated for multiple versions of software and multiple software and applications. If any version vulnerability of any software or application exists on targets network, the cyber exploit will be successful. On the information layer, open source databases e.g. Exploits Database provides categories of exploits based on target point of vulnerabilities. For example, there are remote exploits, web application exploits. DOS & POC exploits and privilege escalation exploits.

SN	EXPLOIT CATEGORY	PERCENTAGE OCCURRENCE
1	Web Application Exploits	58.3%
2	Remote Exploits	17.5%
3	Privilege Escalation Exploits	14.5%
4	DOS & POC Exploits	9.4%

Table 2-2: Exploits Categories *Source (Author)*

2.9 CONCEPT OF EMPLOYMENT

The concept of the employment suggests the tactics, technique, procedures and methods used in achieving the research aims and objectives. The following sections outline and discuss the theories behind the techniques used in implementing the entangled cyberspace framework.

2.9.1 Time Series

Time series analysis assumes an internal structure within data points chronologically recorded over some past time period. Therefore, it deals with understanding the dynamic consequences of events over a period of time. Chronological data are often generated during monitoring, observation or tracking process of a phenomenon or metric. Quantitative time series analysis methods make use of patterns in historical data defined as functions of the observations to infer future values of the variable of interest. A time series recording observations of a single variable is referred to as a univariate time series while observations of more than one variable are referred to as multi-variate. A time series can also be continuous or discrete. A continuous time series has observations measured at every instance in time, whereas a discrete time series has observations recorded at discrete intervals. For example, stock market prices, weather readings can be measured as continuous time series while population growth; product sales can be measured as discrete time series. A time series is typically affected by four major components: the trend, the cyclical component, the seasonal component and the irregular components. Therefore, a time series can be decomposed into its individual components as:

$$Y_t = S_t + T_t + \varepsilon_t$$

Equation 2-6: Time Series Components (Additive Model)

Where Y_t is the time observation at time period t, S_t is the seasonal component at time period t, T_t is the trend component at time period t and ε_t I sthe errors or irregular terms at time period. This model can also be expressed as a multiplicative model as:

$$Y_t = S_t \times T_t \times \varepsilon_t$$

Equation 2-7: Time Series Components (Multiplicative Model)

Multiplicative models are appropriate for non-stationary time series. The multiplicative model can also be derived by taking the logged values of the components in the series:

$$\log(Y_t) = \log(S_t) + \log(T_t) + \log(\varepsilon_t)$$

Equation 2-8: Time Series Components (Logged Model)

A time trend exists when there is a pattern of long-term change in the observed data. A trend may be an increasing or a decreasing trend. For any chronologically observed variable, a time trend is defined as a function of current values, past values and white noise.

$$Y_t = Y_{t-p} + \varepsilon_t$$

Equation 2-9: Time Trend Equation

Where Y is a chronological variable of interest, t is the current time, p is the amount of time into the past (lags) in observation and ε_t is the noise term related to Y_t.
A seasonal pattern exists when a time series in influenced by seasonal factors. For example, a pattern of occurrence every quarter of the year. Seasonality is usually for a fixed or known period.
(Granger, 1980; Sims, 1980; Hamilton, 1994; Luetkepohl, 2011; Enders, 2014) extensively cover methods for fitting time series data for forecasting in research. Time series equations are based on the popular difference equations - an expression relating a single variable Y_t, to lagged values of itself which is an auto-regressive (AR) process with the order of p. Univariate time series are fitted to AR process denoted as:

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + A_t$$

Equation 2-10: Univariate AR Process

Where Y_t are the chronological values for the time series, A_t is the white noise and $\delta = \left(1 - \sum_{i=1}^{p} \phi_i\right) \mu$, with μ as the process mean. An auto regressive model is simply a linear regression model of current values of a series Y_t against past values of itself Y_{t-p}. Analysis of AR models are done using the standard linear least squares techniques (OLS) (Montgomery, Peck and Vining, 2001) representing current values of the time series as the dependent variable and its lagged values as the explanatory variable. The linear model here states that the current values of the time series the sum of its lagged values multiplied by the beta coefficients.

$$Y_t = \sum_{i=0}^{k} \beta_i Y_{t-p} + \varepsilon_t$$

Equation 2-11: Univariate AR Process (OLS Representation)

The beta coefficients β_i, are estimated by minimizing the errors of prediction ε_t. The limitations of Ordinary least square methods lie in its inability to handle collinearity between variables and its linear dimension of analysis.
Another technique for fitting univariate time series models is the Moving Average Model:

$$Y_t = \mu + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \cdots - \theta_q A_{t-q}$$

Equation 2-12: Standard Moving Average Model

where Y_t is the time series, μ is the process mean, A_{t-1} are error terms, and $\theta_1 \ldots \theta_q$ are the parameters of the model. The value q is called the order of the moving average model. A moving average model is conceptually a linear regression of current values of a series Y_t, against the random shocks of prior values of the series. The Moving Average Model assumes that random shocks at each time point in the series are from the same normal distribution with mean (μ) = 0 and a constant standard deviation = (σ^2). Furthermore, Moving Average Models allow for the propagation of errors from past to future values of the series with iterative non-linear fitting procedures as opposed to linear least squares methods.

Box and Jenkins (Box, Jenkins and Reinsel, 2008) popularized the Box-Jenkins model, an approach that combines the Auto-Regressive model and the Moving Average Model, Auto-regressive Moving Average Models (ARMA) processes. The Box-Jenkins model is fitted as:

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + A_t - \theta_1 A_{t-1} - \theta_2 A_{t-2} - \cdots - \theta_q A_{t-q}$$

Equation 2-13: Box-Jenkins ARMA Model Representation

Where the terms and parameters stated in the model have the same meanings as the AR and MA model above.
The Box-Jenkins model assumes stationarity with mean (μ) = 0 and a constant standard deviation = (σ^2). Random shocks account for integration in time series vectors which have a linear or non-linear long-term persistent influence on future values of the series (Enders, 1995). The Box-Jenkins recommendation for getting stationary time vectors from stochastic processes is differencing to the integrated order of the series, producing an ARIMA model (Auto Regressive Integrated Moving Average) (Enders, 2004). A time series sequence with no deterministic component, by which a stationary AR process is derived after differencing d times is said to be integrated order of d denoted as $x_t I(d)$(Granger, 1969). In other words, if the series must be differenced d times to achieve stationarity, it is said to be I(d). A first order differenced time series can be derived by:

$$\Delta Y_t = Y_t - Y_{t-1}$$

Equation 2-14: First-Order Differenced Time Observations

Where Δ denotes a change in the series Y_t. ARIMA models are most commonly used in fitting forecast models due to its flexibility to include both AR and MA models.
Although once limited to the social sciences, time series analytical methods have been lately deployed in other fields ranging from biology to forensics. In cyber security, time series analysis become relevant in problems of anomaly detection (Liu *et al.*, 2015), change point detection (Tartakovsky, Polunchenko and Sokolov, 2013) and Network traffic analysis, (Dua and Du, 2013). Challenges associated with employing time series methods to cyber security solutions are related to the representation of variables in a time series model. For example, a scenario that models a prediction function for network cyber-attacks would have to specify a prior representation for the event of a

cyber-attack. This can be represented as the number of incoming packets to a network (Tartakovsky, Polunchenko and Sokolov, 2013) or as any detected anomaly in network traffic. The representation of the outcome variable is a major determinant of the performance of time series and prediction models (Enders, 2014).
Furthermore, in developing causal models from linked events (Winskel, 2011) or networked graphs (Sofiyanti, Fitmawati and Roza, 2015), researchers use methods that integrate numerous formats of data from various sources. These sources act as predictor variables for a given explanatory variable. However, in the case of cyber security analytics or prediction, the time series model is limited to handling quantitative variables. Therefore, variables would have to be re-interpreted or transformed into quantities to fit into the model.
A multivariate stochastic process is a k-dimensional vector process with a real-value function y: Z × Ω → R K such that, for each fixed t ∈ Z, y(t, ω) is a K–dimensional random vector. The realization is a sequence of vectors yt(ω),t ∈ Z, for a fixed ω. It is a function Z → R K. A multiple time series is assumed to be a finite portion of a realization. The underlying stochastic process is therefore called a data generation process (DGP).

2.9.2 Correlation Analysis

Dependence is measured as any observed relationship, whether causal or not between tow random variables. Correlation is a form of dependence that measures the strength of the linear relationship between two random variables A and B. In correlation analysis, we estimate the correlation coefficient R between two variables which measures the strength of the relationship that exist between the two variables. Additionally, the correlation coefficient ranges from -1 to +1 which also indicates the direction of the linear association. The correlation between two variables can either be positive (above 0) or negative 9below 0). For example, a correlation coefficient of 0.98 indicates a strong positive relationship between the two variables while a correlation coefficient of -0.3 indicates a weak negative relationship. The Pearson's correlation coefficient p (Pallant, 2009), requires the dependence between the two variables to be represented by a linear relationship. Given two random variables X and Y, the Pearson's population correlation coefficient is given as:

$$\rho, X, Y = corr(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}$$

Equation 2-15: Pearson's Population Correlation Coefficient

Where E is the expected value operator, σ_X and σ_Y are the standard deviations of X and Y respectively, μ_X and μ_Y are the expected means of X and Y, and $cov(X, Y) = \frac{\sum_{i=0}^{N}(X - \mu_X)(Y - \mu_Y)}{N-1}$. The variances $\sigma_X \; and \; \sigma_Y$ of X and Y measures the variability or spread of the variable scores around the sample mean. The covariance is therefore a simultaneous measurement of the variability of (X, Y) pairs around the mean of X and the mean of Y.
In addition to the Pearson's correlation measure between two random variables, Charles Spearman provides a nonparametric measure of statistical dependence between rankings of two variables by describing how well two variables are related by using a monotonic function (Spearman, 1904). While the Pearson's correlation coefficient is restricted to the assumption of a linear dependence between the two variables, the Spearman's Rank Coefficient (Spearman, 1904) assess both linear and non-linear relationships between the two variables X and Y.

For a given sample size N of two random variables X and Y, the N raw scores $X_I \; and \; Y_i$ are converted to ranks $rgX_I \; and \; rgY_i$ and the spearman's rank correlation coefficient r_s is computed as:

$$r_s = \rho_{rgX,rgY} = \frac{cov(rg_X, rg_Y)}{\sigma_{rgX}\sigma_{rgY}}$$

Equation 2-16: Spearman's Ranked Correlation Coefficient

Where ρ is the Pearson's correlation coefficient applied to the ranked variables, $cov(rg_X, rg_Y)$ is the covariance of the ranked variables and $\sigma_{rgX} \; and \; \sigma_{rgY}$ are standard deviations of the ranked variables.
The Spearman's rank coefficient is usually defined as the Pearson's correlation coefficient between two ranked variables. Additionally, the Spearman's ranked correlation coefficient is shown to be less sensitive to non-normality in data (Ghasemi and Zahediasl, 2012).
Most correlation measures are not sensitive to the scale on which the variables are measured and are unaffected by any scaling transformation applied to the data. However, most correlation measures are sensitive to the manner in which samples are taken and the size of the sample. For example, correlation measures may show different results for a large biased sample collected from a certain part of the population as opposed to a small unbiased sample collected from a representative population. When measuring dependence between two random variables, it is important to consider the sampling population and the sample size (Engle and Yoo, 1987).

2.9.3 Normality and Parametric Assumptions

Many statistical tests such as regression, t-tests, analysis of variance and most parametric tests assumes normality of data observations i.e the population from which the sample data is drawn are normally distributed. Normality is the behaviour of data that is consistent with the normal behaviour of that population. For example, the changes in the prices of Apple stock prices are, under normal circumstances, expected to follow a long-term equilibrium around the mean at any given point in time. Consistent drifts from the long-rum equilibrium suggests that the trends in share prices are not stable. It has been shown in literature (Ghasemi and Zahediasl, 2012) that when the assumptions of normality do not hold, it is difficult to draw accurate and reliable conclusions from the estimated parameters of analysis.
Theoretically, for each mean and standard deviation combination, a normal distribution based on proportions can be determined. The percentile bell curve for a standard normal distribution is shown below.

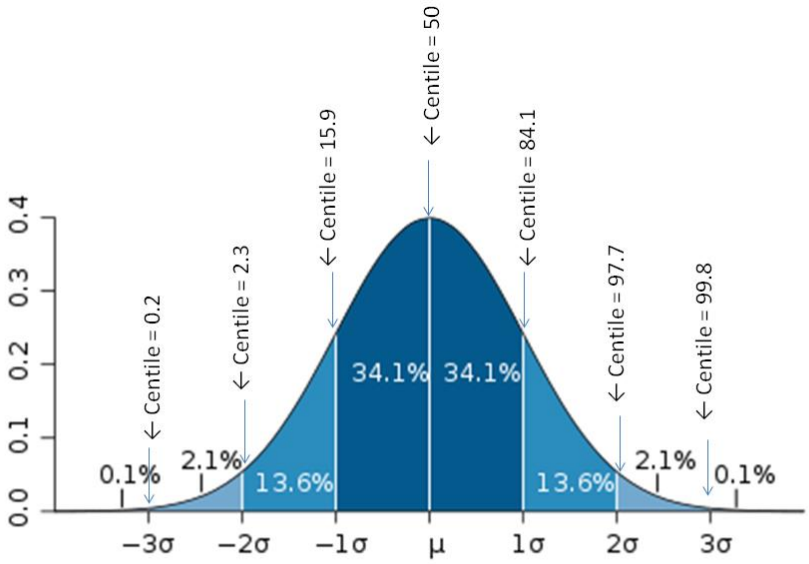

Figure 2-6: Bell Curve For Standard Distribution

A percentile is a measure indicating a value below which a given percentage of observations within the sample data fail to meet certain conditions. For example, the 20th percentile is the point at which 20% of the observations within any given context of analysis, should fall. The graph above shows a normal distribution should have 68% of the data points should falling within 1 standard deviation of the mean or 95% of the data falling within 1 standard deviation of the mean or 99.5% of the data falling within 3 standard deviations of the mean.

There are broadly 2 main methods for determining if a distribution is normal: a) Graphical/Visual methods b) Statistical testing methods. Although, visual inspections of statistical distributions as a basis for establishing normality have been proven to be unreliable, they are a good way of determining the distribution of data. There are several graphical methods for inspecting the distribution of a data: the frequency distribution (histogram), the stem-and-leaf plot, boxplots, cumulative frequency plots (P-P plots), quantile probability plots (Q-Q plots) and the standard bell curve. Visual representation of data simply shows the distribution of data in a sample but gives to formal justification for establishing normality.

Statistical tests for normality are considered more reliable as they put a measure to the degree of normality in a sample data. The normality tests are usually supplementary to visual tests but are more precise since probabilities are calculated. The main tests for normality are Kolmogorov-Smirnov Test (Massey, 1951) , Shapiro-Wilk Test (Shapiro and Wilk, 1965), Anderson-Darling Tests (Anderson and Darling, 1954), Anscombe-Glynn Test (Anscombe and Glynn, 1983), D'Argostino skewness test (D'Agostino and Pearson, 1973). The most commonly used tests in literature however are the Shapiro-Wilk tests, the Kolmogrov-Smirnov test and the Anderson-Darling test. Most tests for normality estimate the probability that the sample observations were drawn from a normal distribution. Theoretically, the tests are conducted to test the differences between two groups of observations to determine if the two data sets are from the same group. The hypothesis tests are set up as follows:

$$H_0: The\ observed\ data\ points\ are\ not\ significantly\ different\ from\ a\ normal\ distribution$$
$$H_1: The\ observed\ data\ points\ are\ significantly\ different\ from\ a\ normal\ distribution.$$

For small sample sizes, most normality tests are meaningless while for large sample sizes, significant results would be returned even in cases of moderate deviations from normality. The Kolmogorov-Smirnov test is an empirical distribution function in which the cumulative distribution function of the test distribution is compared to the empirical distribution of the data. The Shapiro-Wilk test is based on the correlation between the data and an assumed normal distribution score. The Kolmogorov-Smirnov test is criticized for being too sensitive to outliers (Chen and Liu, 1993) and proven to have a lower probability to reject false null hypothesis therefore increasing the probability of Type II errors.

2.9.4 Parametric Assumptions and Data Transformations

Both normality and stationarity are pre-requisite assumptions to various time series analysis techniques. Transformations for normality are performed when the data does not meet the assumptions of normality. There are three main ways to normalize data; a) Mean Normalization b) Logarithmic transformation and c) square-root transformations.

2.9.5 Co-integration Analysis

In time series analysis, there is often the assumption of a long-run relationship between a group of variables of interest. For example, analyst may suspect that the stock prices of apple (AAPL) may move together with the company's sales figures. The idea of co-integration is that there is a common stochastic trend between the pair of variables. The aim of co-integration analysis is the detection and analysis of long-run relationships amongst a set of variables. Furthermore, co-integration analysis acts as a pre-requisite for validity of theoretical assumptions e.g causal relationships.

Given a pair of time series processes $x_t$and y_t, both I(d) where d ≠ 0, if there exist some linear combination of $x_t$and y_t, for example, $y_t - \beta x_t$, that is I(0), then $x_t$and $y_t$are co-integrated with β as a scaling factor. A linear co-integrating relationship exist between two integrated time series $x_t$and y_t if there is some form of long-term true relationship between $x_t$and $y_t$that holds for all t ……. t-n. Here, it is theoretically acceptable for the time series to be non-stationary provided they are both I(d) where d ≠ 0 and $I(d)_x = I(d)_y$. In testing for co-integration between two variables, a linear least squared regression equation is constructed to derive estimates of the generated error term. If X_t & Y_t are non-stationary and co-integrated, then some linear combination of the two-time series vectors must return stationary ε_t. That is:

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

Then;

$$\varepsilon_t = \alpha - \beta x_t - y_t$$

Equation 2-17: Non-Stationary Co-integrated Time Vectors

Where $\varepsilon_t$is stationary with $\mu_t \approx 0$and variance = σ^2. A vector auto-regressive model is an n-equation, n-variable summarised representation of the complex dynamics of a multi-dimensional system (Granger, 1980; Enders, 2014).

Given that most time vectors are often non-stationary, they require differencing or de-trending to be transformed to a stationary state. The problem with differencing or de-trending in co-integration analysis is the risk of removing long-run relevant information. One possible way of proceeding is to represent the models as error correction models (ECM). It has also been shown (Galenko *et al.*, 2009) that if a set of variables are co-integrated, there exist some error correction representation of for the

set of variables. An error correction model is a multi-variate time series analysis technique most used for data where the underlying variables have a long-run and short-run stochastic trend. The model propagates deviations from the long-run equilibrium to the short-run dynamics. Therefore, Error Correction models estimate the speed at which the response variable returns to equilibrium after a deviation caused by changes in the explanatory variables. The error correction representation has an added advantage as it preserves both long-run and short-run relationships.

Sagan (1964) first recommended error correction models as a solution to the problem of spurious regressions put forward by Yule (1936) and Granger & Newbold (1974) in time series analysis. Given two integrated non-stationary time series, there may exist statistically significant relationship between the two variables although these variables may be unrelated. Research provide three main methods for testing for true co-integration in dynamic models.

One of these methods is the Engle and Granger two-step approach. Engle and Granger (Engle *et al.*, 2017) recommend estimating the Error Correction Model in one step and estimating the vector error correction model in a later second step. The first step of the Engle-Granger approach is testing for non-stationarity in the time series of interest. The standard unit root testing and the Augmented Dickey Fuller test (Cheung and La, 1995) determine if the series are stationary and if the errors are linearly correlated. Given two non-stationary time series vectors integrated order 'd', where d > 0, the Engle-Granger approach estimates the Error Correction Model as follows:

$$A(L)\Delta Y_t = \gamma + B(L)\Delta X_t + \alpha(Y_{t-1} - \beta_0 - \beta_1 Y_{t-1}) + \upsilon_t$$

Equation 2-18: Engle-Granger Error Correction Model

The variables are co-integrated by the Engle Granger's representation if this ECM can be produced with the two variables. If the regression as tested in step 1 of the two-step approach is not spurious, the second step is estimating the model using the ordinary least squares methods. The residuals from the prediction equation $\varepsilon_t = Y_t - \beta_0 - \beta_1 X_t$ are propagated to a regression of the first order differences of the variables plus a lagged error term as follows:

$$A(L)\Delta Y_t = \gamma + B(L)\Delta X_t + \alpha\varepsilon_{t-1} + \upsilon_t$$

Equation 2-19: Error Propagation in First-Order Differenced Lagged Time Vectors

The Engle-Granger approach is limited as it allows for a single co-integrating relationship between a pair of variables.

Another method provided for co-integration testing is the Johansen test which allows for testing co-integrating relationships between more than a pair of variables. It is a procedure for testing co-integration between several I(1) variables that permits more than one co-integrating relationship. There are two types of Johannsen's test for co-integration: with trace or he maximum eigen value. The trace method assumes that the number of co-integrating variables is less than the number of variables in the model. The trace tests the alternative hypothesis that the co-integrating rank r is at least {1 …. k}. Where k is the number of variables in the system and r is the co-integrating rank. The maximum eigen value test is the same as the trace test except the threshold for selection of r is r+1 rather than r as with the trace tests. The Johansen's test is limited as it is only applicable to large samples where N is at least 100 as shown by (Johansen, 2000).

The Phillips-Ouliaris co-integration test (Phillips and Ouliaris, 1990) show that applying unit-root tests to co-integrating residuals do not produce the expected distributions for the dickey-fuller test under the null of 'no co-integration'. Phillips and Ouliaris prove that due to plausible spurious

regressions, the distributions of these tests are affected by the number of deterministic trend terms and the number of variables being tested for co-integration.
Co-integration analysis is superior for testing long-term relationships between variables of interest which implies that co-integration techniques are optimized for large sample sizes. Although there are methods that allow for short-run analysis, studies have shown the limitations of such applications.

2.9.5.1 Dimensional Co-Integration

In view of the multi-dimensional cyberspace discussed in section 2.8.4, co-integration may occur within dimensions or between dimensions in cyberspace. The potential of this is referred to in this research as dimensional co-integration. Dimensional co-integration is formally defined as the identification of co-integrating vectors within or between data dimensions in cyberspace. Dimensional co-integration can occur within dimensions of which we refer to as 'Inter-dimensional Co-integration" or between dimensions of which we refer to as "Intra-dimensional Co-integration".

2.9.5.1.1 Intra-Dimensional Co-integration

Given a multi-dimensional dataspace Ω ***with time series vectors representative of incidents in cyberspace across various dimensions of cyber space*** X_t, Y_t ***and*** Z_t***, if the co-integrating rank r for any pairs of variables across*** X_t, Y_t ***and*** Z_t ***is greater than 1, we say these variables are intra-dimensionally co-integrated.*** $r_\Omega > 1$**.**

2.9.5.1.2 Inter-Dimensional Co-integration

Given a multi-dimensional dataspace Ω ***with time series vectors representative of incidents in cyberspace across various dimensions of cyber space*** X_t, Y_t ***and*** Z_t***, if the co-integrating rank r for any pairs of variables within*** X_t, Y_t ***or*** Z_t ***is greater than 1, we say these variables are inter-dimensionally co-integrated.*** $r_{XY}, r_{XZ}, r_{YZ}, r_{X,Y,Z} > 1$

2.9.6 Vector Auto-Regressive Models

Non-linear methods highlight the limitations of linear assumption between series. This limitation necessitated the need A vector auto-regressive model (Sims, 1980) is a generalisation of the univariate autoregressive model for forecasting a collection of variables therefore providing a mechanism to quantitatively summarize the complex dynamics of events in cyberspace on a time spectrum. Vector auto-regressive models provide empirical evidence of the responses of variables to various exogenous impulses. As Baille (1979), Sims (1980), Toda & Yamomoto (1995) and others argue in early influential papers, vector autoregressive models provide a coherent and credible approach to capturing rich dynamics in multiple time series simultaneously. Additionally, feedback relationships (Granger, 1980) are allowed for in VAR.
A Vector Autoregressive model is an n-equation, n-variable representation where each variable is predicted by its own lagged values plus the current and past values of the remaining n-1 variables (Zivot and Wang, 2006; Luetkepohl, 2011). The VARs eliminates the assumptions of a one-directional linear dependency between variables that exist within the conventional AR (p) models. For four variables$y_{1,}y_{2,}y_{3,}y_4$, we construct a 4-variable, 4-equation VAR (1) model to represent the evolution of these four variables over a period of time as explained by their lagged values at 1 step into the past.

$$y_{1t} = c_1 + A_{11}y_{1,t-1} + A_{12}y_{2,t-1} + A_{13}y_{3,t-1} + A_{14}y_{4,t-1} + \varepsilon_{t1}$$
$$y_{2t} = c_2 + A_{21}y_{1,t-1} + A_{22}y_{2,t-1} + A_{23}y_{3,t-1} + A_{24}y_{4,t-1} + \varepsilon_{t2}$$
$$y_{3t} = c_3 + A_{31}y_{1,t-1} + A_{32}y_{2,t-1} + A_{33}y_{3,t-1} + A_{34}y_{4,t-1} + \varepsilon_{t3}$$
$$y_{4t} = c_4 + A_{41}y_{1,t-1} + A_{42}y_{2,t-1} + A_{43}y_{3,t-1} + A_{44}y_{4,t-1} + \varepsilon_{t4}$$

Equation 2-20: 4-Variable 4Equation VAR(1) Model

The matrix form of the equations above is:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} + \begin{bmatrix} A_{11}A_{12}A_{13}A_{14} \\ A_{21}A_{22}A_{23}A_{24} \\ A_{31}A_{32}A_{33}A_{34} \\ A_{41}A_{42}A_{43}A_{44} \end{bmatrix} \begin{bmatrix} y_{t1-1} \\ y_{t2-1} \\ y_{t3-1} \\ y_{t4-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \end{bmatrix}$$

Equation 2-21: Variable 4Equation VAR(1) Model, Matrix Representation

Where $y_{1,}\, y_{2,}\, y_{3,}\, y_{4\ are}$ stationary time-dependent vectors with mean = 0 and variance = σ^2.

Vector auto-regressive models are used for structural analysis of multivariate time series where each variable is a linear function of past lags of other variables. The VAR generates forecasts for each variable in the system, equation by equation using the principle of least squares. The parameters for each variable are therefore estimated by minimizing the sum of squares. Vector auto-regressive models are a unique case of the general VARMA models presented by (Lütkepohl, 2004). VARMA models multivariate time series with a Vector auto-regressive structure along with corresponding moving averages. VAR models eliminate the limitations of a unidirectional relationship common to the generalized OLS time series models.

Vector Autoregressive models often used in analysing the effects of structural shocks to a system are normally subject to a critical determination of the lag length of the model. Braun and Mittnik (1993) demonstrated the inconsistencies of VAR models whose specified lag lengths are different from the true lag length. Lütkepohl (1993) also indicates the association between lag lengths and the mean square errors and auto correlation between variables in the system. Information criteria are statistical methods for measuring the goodness of fit of a model that penalizes for complexities. Various information criteria differ in the form of penalties given to different models. Ivanov & Kilian (2005) demonstrate how the use of information criteria can be applied to the lag length selection of VAR models. The three main information criteria analysed were the Akaike Information criteria, the Bayesian Information Criteria and the Hannan-Quinn information criteria.

Akaike Information Criteria (Akaike, 1989; Cavanaugh & Neath, 2014)

$$\text{AIC} = -2\left(\frac{logL}{T}\right) + \frac{2k}{T}$$

Equation 2-22: Akaike Information Criteria (Akaike, 1989; Cavanaugh & Neath, 2014)

Bayesian Information Criterion or Schwarz Criterion (Konishi & Kitagawa, 1996; German et al, 2014)

$$\text{BIC/SQ} = -2\left(\frac{logL}{T}\right) + \frac{klogT}{T}$$

Equation 2-23: Bayesian Information Criterion or Schwarz Criterion (Konishi & Kitagawa, 1996; German et al, 2014)

Hannan-Quinn criterion (Aznar & Salvador, 2002)

$$\mathrm{HQ} = -2\left(\frac{logL}{T}\right) + 2k\frac{ln(logT)}{T}$$

Equation 2-24: Hannan-Quinn criterion (Aznar & Salvador, 2002)

Although recently, the techniques of structural vector autoregressive models have been used for predictions in random processes such as stock prices Kuo (2016) predictions and crude-oil prices Zhao et al (2015), macro-economic variables Ayadin & Cavdar (2015), its methods remain limited in application in the area of cyber defence. In the context of a multi-dimensional cyberspace, VARs may be applicable to developing a structural approach for representing data from the multi-dimensions of cyberspace. VARs are equally efficient for performing impact-response analysis as a method for identifying cause and effects between variables of interest. This method has been extensively used in the financial sector as in Silignakis (2011), Olson (2014), Jin & An (2015) and have proven to accurately capture the stochastic dynamics in financial markets to a certain degree of certainty.

It also important to understand the challenges that the application of Vector Autoregressive models poses in cyber defence. Data relevant to the prediction of cyber incidents are deposited in various file formats in cyberspace. The VAR model expects as input, stationary time-dependent numeric vectors (Zivot and Wang, 2006). The numeric quantification of the various types of formatted data in cyberspace so that the context of information contained is retained constitutes a challenge to multivariate analysis of data across cyberspace, for example, the measure of 'cyber negativity' or 'cyber positivity' of a sentence or the representation of the details of a cyber-relevant video or image. Methods of transforming cyber data have been concentrated in the areas of natural language processing and analysis of text formatted data (Pang and Lee, 2008; Grimmer and Stewart, 2013; Haddi, Liu and Shi, 2013). The techniques described by these researchers are fundamentally dependent on the presence of a prior human or computer-generated lexicon (O'Connor *et al.*, 2010; Tang *et al.*, 2014) and quantifications are limited to the existing classification of specific words in the lexicon.

Several criticisms of the VAR approach to analysis of multi-dimensional data centre on the relative capacity of the model to handle high-dimensional data (Benanke et al, 2005). The restrictions on the amount of variable to include in a single VAR model may lead firstly, to sacrificing information included to only a selected few, therefore relationships are only captured for variables included in the VAR.

2.9.7 Vector Error Correction Models

The VAR framework is suited to variables that are stationary in the levels i.e. I (0) VARIABLES. If the series are not stationary in the levels, then the framework needs to be modified for consistent estimation of the relationships among series. The Vector Error Correction Model (VECM) is a case of VARs for variables that are stationary in the differences i.e. (I (D) where $D \neq 0$. In a system of multiple evolving I (0) variables, there may exist several independent co-integrating vectors in which linear combinations of these variables are co-integrating. An alternative terminology for the Vector Error Correction Model is the Co-integrated VAR model. The use of co-integration in the Vector Auto regressive framework, consider a data generation process with a K-dimensional VAR (p) process without deterministic trend terms. If the set of K-dimensional variable are I (1), the OLS regression analysis becomes invalid. However, there may be one or more equilibrium relationships in which the co-integrating rank of the K-dimensional vector r, can be estimated. Theoretically, VECMs are representations of co-integrated VAR by Granger's Representation theorem (Luetkepohl, 2011). The VECMs studies how deviations from the long-rum equilibrium are corrected. The Vector Error

Correction Model restricts the long-run behaviour of the endogenous variables to converge to their long-run equilibrium relationships while allowing for short-run dynamics. This is achieved by including an Error Correction Mechanism in the model to correct a proportion of disequilibrium from one period to the next (Engle *et al.*, 2017). The Error Correction mechanism is usually incorporated into VAR in the differences (Hyndman and Koehler, 2006) which further evolves into the Vector Error Correction Model. Given the standard VAR representation:

$$Y_{1,t} = \beta_{10} + \beta_{11}Y_{1,t-1} + \alpha_{11}Y_{2,t-1} + \varepsilon_{1,t}$$
$$Y_{2,t} = \beta_{20} + \beta_{21}Y_{2,t-1} + \alpha_{21}Y_{2,t-1} + \varepsilon_{2,t}$$

Equation 2-25: VECM In the Differences

the VECM is represented as the VAR in the differences, with an added ECM, given by $\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}\left[Y_{1,t-1} - \gamma Y_{2,t-1}\right]$. Provided that $Y_{1,t}\ and\ Y_{2,t}$ are co-integrated with γ as the co-integrating coefficient, $\left[Y_{1,t-1} - \gamma Y_{2,t-1}\right]$ will be stationary with mean = 0 and variance = σ^2. The longrun relationship between the two variables is therefore defined by γ. $\beta_1\ and\ \beta_2$ are called the error correction coefficients and measures the proportion of last period's disequilibrium that should be corrected in the next period.

2.9.8 Text Quantification Techniques

This section describes theories supporting the text mining techniques as used in research to deal with text data. Text mining involves the application of techniques from areas such as information extraction, natural language processing, and information retrieval to quantitatively analyse text data (Jason, 2007). Text data is usually a sequence of unstructured data called documents such that useful information is hidden within a lot of useless information in context. Many text mining techniques have been applied in contemporary research areas such as economics and finance. This research attempts to complement the works of (Tsai and Chan, 2007; Hernández *et al.*, 2016; Lippmann *et al.*, 2016) in applying these text mining techniques to areas of cyber research. Particularly, this research curates a number of these techniques with an emphasis on methods for quantifying text data for further numerical analysis. The following sections presents literature on various methods for text quantification.

2.9.8.1 Frequency and dictionary-based techniques

In contrast to complex quantitative methods found in contemporary research, (Blumenstock, 2008) has shown that a simply metric, the number of words in a text corpus is a reliable measure of information quality. (Blumenstock, 2008) shows that a simple measure of "text length" performs better in classifying featured articles versus random articles on Wikipedia. The word count is a derivative of the popular bag-of-words model an information retrieval technique used in natural language processing. The bag-of-words model is sometimes called the vector space model. This technique represents text data, such as a sentence or a document as a multiset of its words, with no reference to grammar, context or similarity. Simple derivatives of the bag-of-words model such as word frequency algorithm show the number of words in a document or text. In contrast to complex

quantitative methods found in contemporary research, (Blumenstock, 2008) has shown that a simply metric, the number of words in a text corpus is a reliable measure of information quality.

2.9.8.2 Opinion mining (Sentiment)

In the last 3 decades, quantifying "how people feel" has become a new trend in an era of free and open source information. Opinion mining refers to techniques used to systematically identify, extract and quantify affective states and subjective information from what people say. Sentiment analysis seeks to extract 'opinions" from human-generated text. This techniques adds to a body of research focused on processing and analysing information generated on the social dimension of cyberspace (Wilson, Wiebe and Hoffmann, 2005; Pang and Lee, 2006). Previous methods of obtaining such information relied heavily on questionnaires, surveys and interviews which bear the risk of inaccuracy. However, with the dawn of social platforms, information informally and willingly generated by users tend to capture deeper feelings as would be captured by traditional survey methods. An opinionated sentence is a sentence that expresses explicit or implicit positive or negative opinions. Quantitative measures of opinion include subjectivity, polarity, and sentiment. The sentiment of a text measures the degree or strength of its negativity or positivity as measured by its orientation (Pang and Lee, 2006).
Furthermore, (Sproat, 2000) defines sentiment as the underlying felling, attitude or emotion associated with an opinion. This definition provides four main components of an opinion: the opinion holder, the opinion target, the sentiment of the opinion and the time of the opinion
The basic task of most sentiment analysis is polarity classification or quantification. Sentiment analysis seeks to assign a value to a given text corpus measuring the degree of negativity, positivity or neutrality of the text corpus. A document's orientation or polarity is a measure of whether a subjective text expresses a positive or a negative bias on its subject matter. Subjectivity refers to deciding whether a text as a factual nature in terms of accurately describing a given situation without expressing bias (Lui, 2015).

2.9.8.3 Entropy

Considering the information revolution, information represents one of the most important resources of human and societal development. Information and entropy characterize a system from the point of view of both order and uncertainty, in as such that if information measures truth and certainty, entropy measures chaos and uncertainty. Formally, information entropy is the sum of the negative logarithm of the probability mass function of each data value in a piece of information. Information entropy can also be seen as a measure of information randomness and uncertainty of truth in a piece of information or simple the quantitative measure of disorder in a system. There are three variants of entropy in the present day: in thermodynamics according to Clausius, in molecular systems as a measure of disorder, randomness or uniformity and in the theory of information according to Shannon (Claude E Shannon, 1948) as a numerical measure of credibility or truth of information transmitted through an information channel. Although these three variants of entropy have become important in their various fields, Shannon's theory of communication has been most applied in areas if information sciences and text mining (Berger, Pietra and Pietra, 1996; Nigam, Lafferty and Mccallum, 1999; Scholz and Conrad, 2013; David and Thomas, 2015).
In many information science tasks, the subject of study is a given language text with which a mathematical model for analysing the text must be formed. According to Shannon (Claude E

Shannon, 1948), this model is built to preserve the sum of truth and uncertainty, thereby creating a model of text perfection. Shannon (Claude E Shannon, 1948) provides a mathematical formula for defining the entropy of a text:

$$H = -\sum_{i=1}^{N} p_i log_2 p_i$$

Equation 2-26: Shannon's Entropy

Where p_i is the probability of detecting any system unit in their multitude $\sum_{i=1}^{N} p_i = 1, p_i \geq 0, i = 1,2,3..N$

2.9.8.4 Emotion Detection

Despite the wide applications of sentiment analysis in modern text mining, none of its methods supports the recognition of various emotions from the given opinion expressed in text format. Given some basic text, emotion recognition tasks seek to detect what emotions the writer expresses in the opinion of the text. In order to achieve this goal, emotion recognition algorithms create a basis for text labelling and classification based on certain expressions present in the text. Therefore, emotion detection from text may be seen as a multi-class classification problem, to which multiple algorithms are available in research. In the area of emotion classification, there are three main body of works currently popular in research: Profile of Mood States (Norcross, Guadagnoli and Prochaska, 1984), Paul Ekman's six-dimensional classification (Ekman, 1992) and Robert Plutchniks eight-dimensional pairwise emotion classification (Plutchik, 1982).

POMS is a psychological instrument that defines a six-dimensional mood state representation (Norcross, Guadagnoli and Prochaska, 1984). Profile of Mood States (POMS) defines 65 adjectives rated by the opinion expressed on a five-point ranking scale. Each adjective contributes to one of the six categories. For example, feeling tired will positively contribute to the fatigue category. The higher the five-point rank for the adjective, the more it contributes to the overall score for its category. POMS combine these ratings into a six-dimensional mood state representation consisting of categories: anger, depression, fatigue, vigour, tension and confusion. Each dimension is therefore characterised by the presence of a set of emotional adjectives and each individual feeling is accessed by the intensity of the opinion expressed.

Pluchnik's wheel of emotions research shows that there is a total of 34,000 unique emotions that can be used to access various intensities of human feelings. However, these 34,000 emotions were reduced to only 8 primary emotion dimensions: anger, fear, anticipation, surprise, joy, sadness, trust and disgust. Robert Plutchnik extends Ekman's classification and defines a wheel like diagram with a set of eight basic, pairwise contrasting emotions: joy and sadness, trust and disgust, fear and anger, surprise and anticipation.

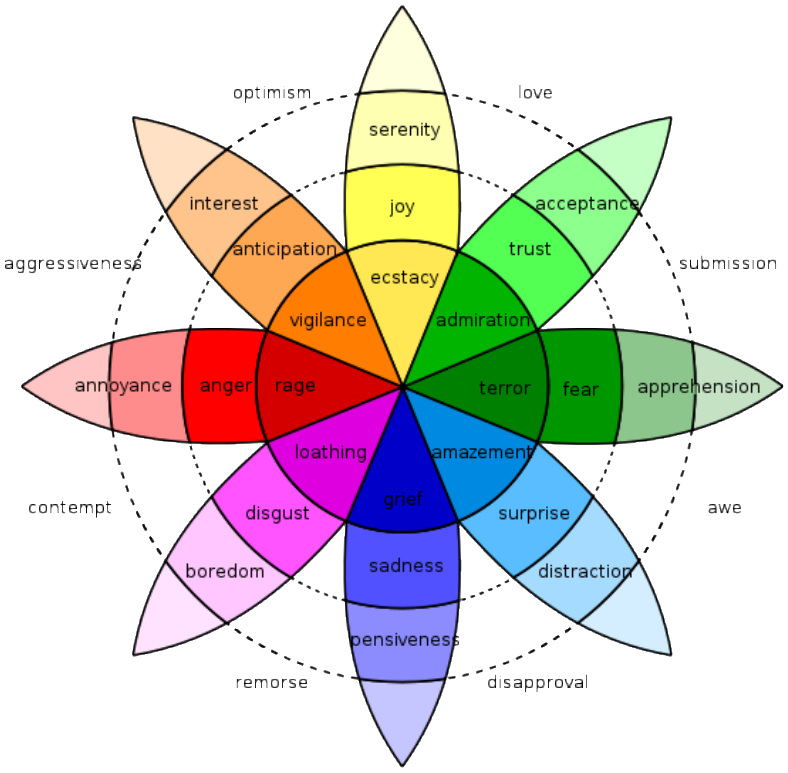

Figure 2-7: Pluctchik's Wheel Of Emotion Classification

Colneric & Demsar (Colneric and Demsar, 2018) approach the problem of emotion detection with by combining the concepts trained recurrent neural networks to labelled data of 73 billion tweets based on Plutchnik's eight dimensions of emotions and Ekman's six-dimensional classifications of emotions to predict emotions from text. Trained recurrent neural network are a set of popular sequence models usually employed in natural language processing tasks. In traditional neural networks, it is assumed that all inputs and outputs are independent of each other (Bishop, 2006). For example, it is assumed that the previous word in a sentence is independent of the next word in the sentence. RNNs are recurrent because they perform the same function for every element of a sequence, with the output being dependent on the previous computations in the sequence. Previous studies in emotion detection and classification were based on Profile of Mood States (POMS). Plutchnik's classification is an expansion of the POMS six-dimensional mood state and the Ekman's six-dimensional emotions.
Researchers have criticized RNNs as being difficult to train and do not yield reliable results when applied to new data in a new linguistic context. Additionally, RNNs are limited in their ability to look back in the sequence beyond a few steps backwards.

2.9.8.5 Lexical Analysis (Context Relatedness)

Context relatedness measures a text's similarity to a given context of analytical interest. For example, calculating how similar a text is to cyber-related activities requires measuring the amount of words in the text that are related to cyber activities. This technique is following a dictionary method which operates on a document-matrix DM derived from the text document. There are two main steps involved in matching a given text to a context of analytical interest. The first step defines a list of keywords or words that capture the context of interest. For example, let a dictionary for 'flu-related'

complaints be DM = {vomit, ache, pain, sneeze, cold, fatigue}. One could then represent the context relatedness (as it relates to a flu incident) of each document in D as: #vomit + #ache + #pain + #sneeze + #cold + #fatigue. The second step represents each document in terms of the normalized frequency of words in the dictionary. The normalization step is necessary as it considers the amount of information generated within the context of analysis. I.e. the ratio of context related words to the amount of words used in document. Similar methods have been applied in natural language processing tasks such as emotion detection (Calefato, Lanubile and Novielli, 2017) where specific adjectives are rated by subject. The presence and intensity of certain adjectives in the given text simply classifies the text under related subjects. These methods are only as effective as the wordlist of adjectives used in the classification task. Wordlists that are not robust enough may lead to under performance of classification model while wordlists that are too robust may lead to models that misclassify texts.

The goal of building a lexicon is to extract a set of terms that capture the context of lingual analytical interest. From an existing sample of 'domain-related' texts, we hope to extract a set of terms that are seen to appear frequently and have higher degree of importance in these texts. Typically, when building domain-specific lexicons, two design approaches are considered and most often combined: selecting and categorising terms within some predefined categories (Kipper *et al.*, 2006; Society, 2016) and weighting terms based on the domain of analysis (Baccianella, Esuli and Sebastiani, 2010) and usefulness of terms in a document corpus to the topic of interest (Jurafsky and Martin, 2017).

The first step in most lexical creation tasks is the candidate terms generation. This step involves creating a wordlist or dictionary of terms that are known to appear frequently in discussions of interest. Keyword-based extraction of terms is most common in this step and has been instrumental to creating traditional lexicons (Ntoulas, Pzerfos and Cho, 2005; Olteanu *et al.*, 2014) for natural language and information retrieval tasks such as sentiment analysis (Baccianella, Esuli and Sebastiani, 2010). For example, (Nielsen, 2011) creates a wordlist of positive and negative words for ranking text documents on a scaled range of -1 to 1 (-1 indicating a strongly negative text document and +1 indicating a strongly positive text document). (Allahyari *et al.*, 2017) applies a similar simplistic approach to a multi-class text classification task. Similarly, (Rose *et al.*, 2010) keyword-based approach provides context to the terms used for lexical analysis by creating a network of lexical-semantic relations between words in a document corpus where the meaning of each term in the lexicon is defined only within the context of its relationship with other terms.

Simple keyword-based candidate term selection approach is further extended to include methods for term scoring. For each term that makes it into the candidate set, scoring techniques evaluate the relative importance of that term in relation to other terms in the candidate set (Kaji and Kitsuregawa, 2007) or to a set of pre-defined domain context terms. For example, quantifying the relationship between the terms 'vulnerability' and 'malware' in a document set. Popularly in research, terms are numerically scored by two main techniques: frequency-based scoring techniques (Blumenstock, 2008; Rose *et al.*, 2010; Zhang, Jin and Zhou, 2010) and scores based on associative dependence (Debole and Sebastiani, 2003), (Khan, Qamar and Bashir, 2016), (Jurafsky and Martin, 2017).

However, since term scoring is based on semantic relationships between words in a text document (independent of any externally perceived meanings), most of these techniques are highly sensitive to the domain of analysis and the sentence in which they occur. For example, consider these two phrases that belong to the same text corpora taken from two different sub-reddits0: 'MySQL Database developer needed urgently for a 2-month project in Belfast.' and 'New MySQL database vulnerability

found on Windows operating system. Yet again!!' Assuming a single domain of analysis, the term 'database' in the midst of other domain-related terms such as: ['vulnerability'], ['operation' and 'system'], ['MySql'], should have a higher-ranking score than the same word in the previous phrase.

In this thesis, we aim to develop a generalizable lexicon that can be applied on most social platforms to automatically detect cyber-related messages using a standard set of discussions in various cyber-related online contexts. Our data is representative of these cyber-related contexts and characterize the occurrence of cyber discussions on a variety of social platform types.

2.10 LIMITATIONS OF PREVIOUS STUDIES

The literature above covers studies in cyberspace characterization, the cyber-attack kill-chain, time signal analysis and inter-connectedness on cyberspace. The literature reveals a gap in the availability of an integrated approach to simultaneous multi-dimensional analysis of analysis of events that may be linked to cyber-attacks. Most scholars have focused on a mono-dimension analysis or incorporated evidence has been limited to the network layer of the physical dimension. Some researchers (Ning *et al.*, 2015), have hinted the existence of inter-connected relationships between events in cyberspace while other researchers (Gandhi *et al.*, 2011a), have put forward arguments for a cultural, political and economic dimension to cyber-attacks. These ideas have always been treated as individual approaches to developing sustainable proactive cyber defence strategies. Literature also identifies a gap in the type of proactive defence strategies being implemented in industry. It shows that most of these so-called proactive approaches are implemented after the fact i.e. optimized to deal with defending and mitigating against cyber-attacks. The is a glaring lack of methods that help put defenders ahead of their adversaries. To this end, this research seeks to implement a framework that integrates these multiple approaches with the aim of developing a multi-dimensional framework capable of simultaneously analysing various events in cyberspace.

3 CHAPTER 3: HYPOTHESIZED MODEL AND APPROACH AND THE THEORETICAL DEVELOPMENT OF THE ENTANGLED CYBERSPACE

3.1 INTRODUCTION

This chapter gives an account of the theoretical and conceptual foundations of this research. It starts by briefly discussing the relevance of theory to information systems (IS) research. This chapter critically discusses how the entangled cyberspace framework is developed using a combination of existing theories discussed in the literature. Thus, it addresses how the existing theories/models are used to ground the entangled cyberspace framework. The entangled cyberspace is based on three theoretical foundations: (a) A multi-dimensional cyberspace which is a translation of a perceived characterization of cyberspace (United States: US Army, 2010; Klimburg, 2011; Barnett, Smith and Whittington, 2014), (b) the cyber-attack kill-chain for conceptualizing activities involved in the propagation of cyber-attacks (Hutchins, 2011; Yadav and Rao, 2015) and (c) Vector autoregressive models for multiple time-based event analysis (Sims, 1980; Luetkepohl, 2011).

Furthermore, this chapter gives a clear account of how the combination of these models converges to address the issues the entangled cyberspace framework attempts to solve in the area of proactive cyber defence. The framework is an integration of strategic and operational concepts in alignment with the capacity for technical analysis. Additionally, the hypothesis to be tested with the experimental data is also critically discussed. In conclusion, this chapter delivers a critical analysis of how the existing frameworks and models come together to deliver the research objectives.

Following a critical review of contemporary literature in subject areas of interest, this section develops the theoretical framework as a foundation on which the concepts put forward in this research are tested and evaluated.
The overview of the literature review produces the following conceptual points underpinning this research:

i. The various characteristic uses of cyberspace creates a stratification of cyberspace based on the roles individual components of cyberspace play in modern existence.
ii. These divisions create a multi-dimensional cyberspace, with each dimension acting independently as well as part of a whole.
iii. Cyber-attacks are orchestrated following the phases of the traditional kill-chain model developed by Lockheed Martin (Hutchins, 2011).
iv. The phases in the kill-chain may vary or evolve for different scenarios of cyber-attacks (Konikoff, Harris and Petersen, 2013; Yadav and Rao, 2015).
v. There are indicators of various phases of the kill-chain beyond the physical dimension of cyberspace (Ning *et al.*, 2015).
vi. Events that characterise various phases of the kill-chain occur across multiple dimensions of cyberspace (Clark, 2010; Klimburg, 2011).
vii. Finally, the inter-dependence between these events characterise the nature of entanglements in cyberspace.

The current state of interdependence consolidates the argument in favour of a proactive approach to effectively capture the nature and usefulness of such inter-related network of events. Therefore, the finding from the literature review demands the formulation of a new model or system that would

complement existing reactive and defensive approaches to keep defenders ahead in the game. This is where the entangled cyberspace fits in, where its purpose within the proposed context is as follows:

i. The Entangled Cyberspace will integrate data across multiple dimensions of cyberspace.
ii. The Entangled Cyberspace will use structural multivariate analytic techniques to provide real-time analysis of the interconnectedness between these dimensions of cyberspace.
iii. The Entangled Cyberspace will provide early warning indicators of a given type of cyber-incident.
iv. The Entangled Cyberspace will provide indicators of cyber-kill chain events present across the various dimensions of cyberspace.
v. The Entangled Cyberspace will provide useful active intelligence for noticing early signs of cyber-incidents.

The sole purpose of the entangled cyberspace framework is to capture the nature of entanglements in cyberspace in such a way that it pre-empts cyber-incidents. The cutting-edge aspect of the entangled cyberspace would be the dynamic structural causal analysis of events across multiple dimensions of cyberspace. The tasks above highlight the role of integration, analysis and intelligence of multiple sources of evidence, which in turn helps provide a scaled down summary of the theoretical model as follows:

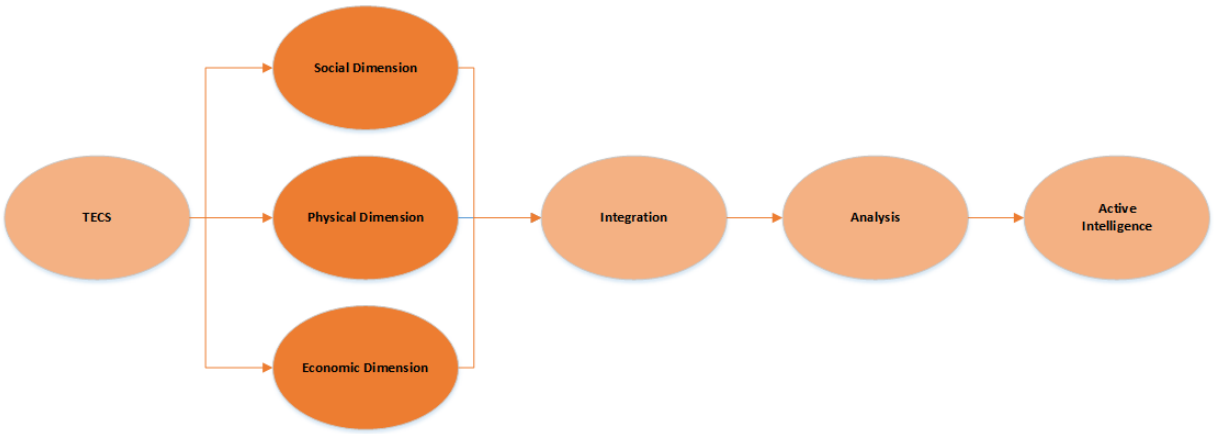

Figure 3-1: Compact Conceptual TECS Framework

This research has argued that the roles of integration and intelligence are extremely crucial for the creation of the entangled cyberspace (TECS) model, which needs to be continually updated with new evidence from events in cyberspace that are predictive of subsequent stages of the kill-chain. The availability of real-time information will form the basis for dynamic testing and selection of predictive features. Unfortunately, there is little literature covering any specific methods for dynamic selection of features in a cross-dimensional, multi-phased structural model. Hahn, Thomas, Lozano and Cardenas (Hahn *et al.*, 2015) went as far as providing a multi-layered framework for analysing the kill-chain in cyber-physical systems. The framework incorporates two main elements:

a) A logical system reference architecture that expresses the architectural composition of multi-layered cyber systems,
b) A multi-layered kill-chain for analysing attacks and threats in multi-dimensional domains.

Thus, the success of the above combination framework prompted this study to rely on features that co-integrate on a cross-dimensional domain in the context of a cyber-attack execution. This research also

employs some cyber forensic techniques, which de-constructs cyber incidents with the aim of understanding the kill-chain, identifying attackers, attackers' motives, means and tactic. Altogether, the literature prompted this study to believe that cyber defenders should direct their intelligence operations towards predictive operations and gathering information that keeps them one step ahead of the attackers.

The entangled cyberspace refers to a seamless integration of multiple sources of evidence across the various dimensions of cyberspace, which characterises the different phases of the cyber-kill-chain on a time spectrum.

The implementation of the entangled cyberspace framework requires multiple sources of evidence be integrated and tested in a structural model. To achieve this, dimensional data observations collected across the dimensions of cyberspace are structurally represented in a VAR/VECM model. Data observations are assumed to measure activities that characterize different phases of the kill-chain. The phases included in each entangled cyberspace model are dependent on the cyber-attack scenario under consideration. Although the model is built around a generic cyber-attack kill chain (Hutchins, 2011), different variations of the model may exist as variations in cyber-attack phases affects the evolution of TECS. The entangled cyberspace assumes a linear dependence between identified features in a multi-dimensional cyberspace. The identified features are further assumed to be predictive of the event of a cyber-attack on the network layer of cyberspace.

3.2 MODEL FOR STRUCTURAL INTEGRATION OF MULTIPLE SOURCES OF EVIDENCE IN CYBERSPACE

In this section, the researcher proposes a complex structural approach to identifying sources of evidence from a multi-dimensional cyberspace for proactive cyber defence. To do this, the researcher characterises cyberspace based on the multi-dimensional framework discussed in this research and considers data generated from these dimensions as information sources for proactive cyber defence models. The rationale behind this approach arises from the need for a multi-dimensional perspective to cyber defence techniques. Moreover, the integration of multiple sources of information into a single model provides a means to study the complex dynamics of events and the nature of entanglements across the various layers of cyberspace.

3.2.1 Data from The Social Dimension

The social dimension is a complex interactive structure made up of individuals, organisations and entities connected by one or more specific types of inter-dependencies such as friendships, interests, likes, dislikes, relationships, political, economic and religious affiliations, relationships of beliefs, kinships, etc. Widely accessible and open source social technologies are increasingly being used by humans to document thoughts, perceptions, opinions and beliefs about real-world events in real-time. Social media platforms like Facebook and Twitter are a rich source of data for proactively countering cyber-attacks. For example, to encourage online user privacy awareness, (Debatin *et al.*, 2009) uses Facebook data to measure users' online exposure to vulnerabilities by studying users' ritualisation, routines, gossips, rumours, this study could examine the relationship of Facebook privacy issues with invasion of privacy. Boub-Harb (Bou-Harb, Debbabi and Assi, 2015) also crawled cyber-crime related messages from Twitter and context-related blog articles to identify notorious hacker groups and their related twitter feeds, blogs and blog activities.

Additionally, (Lippmann *et al.*, 2016) also apply machine learning techniques to cyber-related discussions form social platforms such as Reddit, stack exchange and Twitter to automatically

identify malicious discussions online. Similarly, Burnap, (Burnap and Williams, 2016) applies similar machine learning techniques to Twitter data to automatically identify hateful text directed at minority groups online. Hernández (Hernández *et al.*, 2016) employs user sentiment analysis on a daily collection of tweets in two different contexts; Twitter as a platform for expression of user views and Twitter as a platform to present content related defence threats on the web. Statistical analysis of these data is used to predict the possibility of a future attack.

Data on the social layer of cyberspace exist as a series of interconnected events that captures the real-world personas as well as the cyber personas of online users. Consequently, data and information from social media platforms such as Facebook, Twitter and other micro-blogging platforms that allow users to freely share views, perceptions, and opinions about related matters, feeds into cyber-attack counter models. One distinct advantage of using data from the social layer is the free and open accessibility it provides. However, many researchers have identified data from social platforms as highly subjective and opinionated.

3.2.2 Data from The Physical Dimension

The physical dimension is a bit less complex than the social layer but acts as a gateway to all cyber-attacks as over 90% of all cyber-attacks eventually takes place on the network layer of cyberspace. Data analysis on the physical dimension, most notably the network layer seems to have dominated literature in the past decade. Various types of data that are extracted from the network layer of cyberspace have extensive coverage in literature. Time series models are commonly applied to network traffic data to predict network flow (Barford *et al.*, 2002; Kim and Reddy, 2008), to spot denial of service attacks (Kim and Reddy, 2008), in building intrusion detection systems (García-Teodoro *et al.*, 2009) and classifying network traffic packets (Mahoney, 2003a). Moreover, log data is another source of data for proactive cyber defence models to network traffic flow data. For example, probabilistic latent semantic analysis (Tsai and Chan, 2007) detects cyber defence threats in weblog data files. End-user behaviour can also be inferred using event logs for evidence gathering (Singh and Roy, 2010) and anomaly detection techniques (Patcha and Park, 2007).

Similarly, the use of classification and correlation (Gabra, 2014) techniques on IDS alert logs support cyber forensics and real-time cyber threat assessments. Consequently, data and information from sensors such as the Internet of things (Atzori *et al.*, 2012) also fall within the physical layer of cyberspace. Data on the physical layer may also include some real-world attributes, e.g., weather data, GPS tracking data and some traffic data to trace events such as cyber terrorism or cyber activism (Whiting *et al.*, 2015).

Unlike the social layer, one major disadvantage of using data on the physical layer is accessibility. Consequently, attempts to perform analysis on this layer has always been restricted to context analysis with limited possibilities of generalisation and extension (Baggili and Breitinger, 2015).

3.2.3 Data from The Economic Dimension

The complexity and the interconnectedness of real-world events and cyber events uniquely characterize the economic dimension. Similarly, time series models are also equally applied to stock market data as generated in real-time. Although there is little research on the effects of stock market

prices on cyber defence events, research has covered the integration of stock market data with data from other layers of cyberspace (Bollen and Mao, 2011).

Sufficient evidence supports the extraction of cyber threat intelligence from a systematic analysis of unstructured open source information, e.g., text/news articles (Tsai and Chan, 2007). Similarly, studying past social events in news media and publications reveal factors in the social and cultural dimensions that act as antecedents to cyber-attacks (Whiting *et al.*, 2015). The usefulness of financial prediction markets in information defence risk management is another example this concept with methods such as (Pandey and Snekkenes, 2014) suggests. This research uses this dimension to practically demonstrate claims made by (Gandhi *et al.*, 2011a) of a link between cyber-events and political, cultural, and socio-economic events.

Data representation is the transformation of real-world experiences and perceptions into a computational domain that should provide useful information. Data integration is the unique problem of combining data from different sources on a common attribute in such a way that it retains valuable information from these various sources. The integration method uses a vector auto-regressive structural framework for analysing the evolution of multiple time series simultaneously.

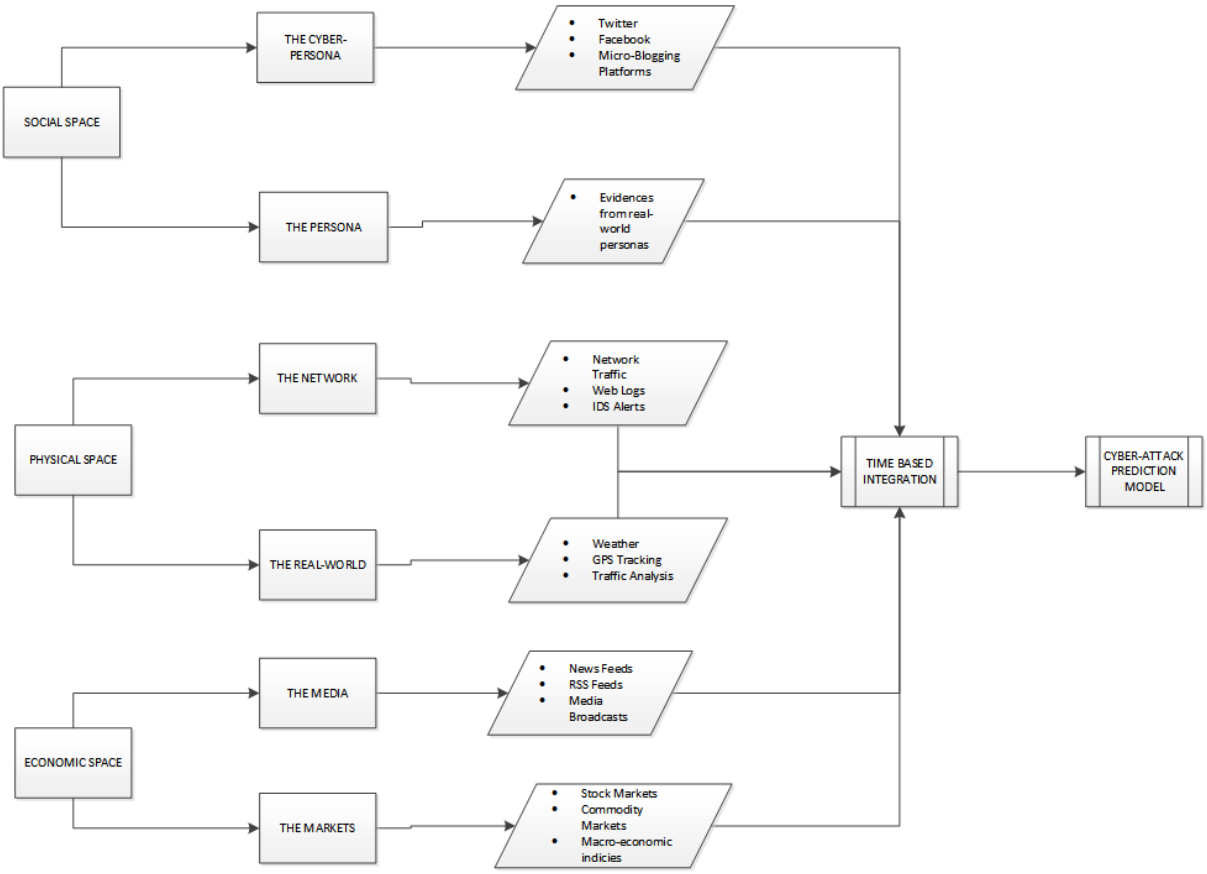

Figure 3-2: First Conceptual Framework for the Entangled Cyberspace (Source: Author)

3.3 IMPLEMENTATION OF ENTANGLED CYBERSPACE THEORY

Following the theoretical development of the entangled cyberspace theory, this section presents the techniques, methods and models for a practical implementation of the developed concepts. This section presents the structural models with which evidence from the various dimensions of cyberspace

would be fitted and evaluated. The researcher constructs a vector autoregressive model for each identified dimension of cyberspace and presents the hypotheses for testing the theories within the developed structural models.

Given a multi-dimensional cyberspace with J identified dimensions (in this study the researcher uses a 3-dimensional cyberspace, J=3), we represent information from these dimensions as Ω_1 *through* Ω_j and construct a basic VAR model for each dimension as:

The physical Dimension

$$\Omega_1 = [x_1 \dots x_n] = \mathrm{X_t} = \begin{cases} \mathrm{x_1} = \mathrm{g_{x,11}x_{1,t-1}} + \cdots + \ \dots \ + \mathrm{g_{x,1n}x_{n,t-1}} + \varepsilon_{\mathrm{x,1t}} \\ \mathrm{x_2} = \mathrm{g_{x,21}x_{1,t-1}} + \cdots + \ \dots \ + \mathrm{g_{x,2n}x_{n,t-1}} + \varepsilon_{\mathrm{x,2t}} \\ \vdots \\ \mathrm{x_n} = \mathrm{g_{x,n1}x_{1,t-1}} + \cdots + \cdots + \mathrm{g_{x,nn}x_{n,t-1}} + \varepsilon_{\mathrm{x,nt}} \end{cases}$$

Equation 3-1: VAR(p) Representation of the Physical Dimension

The Social Dimension

$$\Omega_2 = [y \dots y_n] = \mathbf{Y_t} = \begin{cases} \mathbf{y_1} = \mathbf{g_{y,11}y_{1,t-1}} + \cdots + \ \dots \ + \mathbf{g_{y,1n}y_{n,t-1}} + \boldsymbol{\varepsilon}_{\mathbf{y,1t}} \\ \mathbf{y_2} = \mathbf{g_{y,21}y_{1,t-1}} + \cdots + \ \dots \ + \mathbf{g_{y,2n}y_{n,t-1}} + \boldsymbol{\varepsilon}_{\mathbf{y,2t}} \\ \vdots \\ \mathbf{y_n} = \mathbf{g_{y,n1}y_{1,t-1}} + \cdots + \cdots + \mathbf{g_{y,nn}y_{n,t-1}} + \boldsymbol{\varepsilon}_{\mathbf{y,nt}} \end{cases}$$

Equation 3-2: VAR(p) Representation of the Social Dimension

The Economic Dimension

$$\Omega_3 = [z_1 \dots z_n] = \mathbf{Z_t} = \begin{cases} \mathbf{z_1} = \mathbf{g_{z,11}z_{1,t-1}} + \cdots + \ \dots \ + \mathbf{g_{z,1n}z_{n,t-1}} + \boldsymbol{\varepsilon}_{\mathbf{z,1t}} \\ \mathbf{z_2} = \mathbf{g_{z,21}z_{1,t-1}} + \cdots + \ \dots \ + \mathbf{g_{z,2n}z_{n,t-1}} + \boldsymbol{\varepsilon}_{\mathbf{z,2t}} \\ \vdots \\ \mathbf{z_n} = \mathbf{g_{z,n1}z_{1,t-1}} + \cdots + \cdots + \mathbf{g_{z,nn}z_{n,t-1}} + \boldsymbol{\varepsilon}_{\mathbf{z,nt}} \end{cases}$$

Equation 3-3: VAR(p) Representation of the Economic Dimension

Where $\boldsymbol{X_t}$ refers to a set of time series observations from the physical dimension, $\boldsymbol{Y_t}$ represents a set of time series observations from the social dimension and $\boldsymbol{Z_t}$ represents a set of time series observations from the economic dimension of cyberspace. Therefore, $\boldsymbol{X_t} = \begin{pmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{pmatrix}$, $\boldsymbol{Y_t} = \begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{nt} \end{pmatrix}$ and $\boldsymbol{Z_t} = \begin{pmatrix} z_{1t} \\ z_{2t} \\ \vdots \\ z_{nt} \end{pmatrix}$ are matrices of time series auto-regressive processes, $\boldsymbol{G_X} = \begin{pmatrix} g_{x,11} & g_{x,12} & \cdots & g_{x,1n} \\ g_{x,21} & g_{x,22} & \cdots & g_{x,2n} \\ \vdots & \vdots & & \vdots \\ g_{x,n1} & g_{x,n2} & \cdots & g_{x,nn} \end{pmatrix}$, $\boldsymbol{G_Y} = \begin{pmatrix} g_{y,11} & g_{y,12} & \cdots & g_{y,1n} \\ g_{y,21} & g_{y,22} & \cdots & g_{y,2n} \\ \vdots & \vdots & & \vdots \\ g_{y,n1} & g_{y,n2} & \cdots & g_{y,nn} \end{pmatrix}$ and $\boldsymbol{G_Z} = \begin{pmatrix} g_{z,11} & g_{z,12} & \cdots & g_{z,1n} \\ g_{z,21} & g_{z,22} & \cdots & g_{z,2n} \\ \vdots & \vdots & & \vdots \\ g_{z,n1} & g_{z,n2} & \cdots & g_{z,nn} \end{pmatrix}$ are matrices of auto-regression

co-efficient and $\boldsymbol{\varepsilon_{Xt}} = \begin{pmatrix} \varepsilon_{x,1t} \\ \varepsilon_{x,2t} \\ \vdots \\ \varepsilon_{x,nt} \end{pmatrix}, \boldsymbol{\varepsilon_{Yt}} = \begin{pmatrix} \varepsilon_{y,1t} \\ \varepsilon_{y,2t} \\ \vdots \\ \varepsilon_{y,nt} \end{pmatrix}$ and $\boldsymbol{\varepsilon_{Zt}} = \begin{pmatrix} \varepsilon_{z,1t} \\ \varepsilon_{z,2t} \\ \vdots \\ \varepsilon_{z,nt} \end{pmatrix}$ are matrices accumulating the corresponding residuals which are uncorrelated through time.

The equations above are a representation of the first theoretical model presented in Figure 3-2. The first hypothesis seeks to establish a link between sources of evidence within each dimension.

3.3.1 Hypothesis One

Given a set of multi-variate time series, observations of a set of variables representative of events on the physical, social or economic dimension of cyberspace, Xt, Yt and Zt, the inter-dimensional co-integrating rank of all Xt, Yt and Zt is greater than or equal to 1. i.e. r >= 1.

To extend the basic conceptual framework, the researcher integrates two independent frameworks: the traditional kill-chain model (Hutchins, 2011) and the multi-dimensional cyberspace (Clark, 2010). Given a cyber-attack incident with K phases in its kill-chain, we represent events in a K-phased cyber kill-chain as:

$$E_1|\dots\dots|E_{K-1}|E_K.$$

Equation 3-4: Phased Representation of the Cyber-Attack Kill-Chain

The events at each stage of the kill-chain are hypothesized to occur on different dimensions of cyberspace. For example, a port or vulnerability scan of a target network is spotted on the network layer – the physical dimension of cyberspace. Therefore, with cyber-attack events occurring across a multi-dimensional cyberspace with J=3 dimensions, the researcher hypothesizes a multi-dimensional kill-chain that occurs across the various dimensions of cyberspace. Therefore, we can produce a cross-dimensional multi-phased cyber kill-chain model as a multi-dimensional matrix of data vectors as follows:

	Antecedents	Recon	Weaponize	Delivery	Exploit	Install	C&C	Action
Physical Dimension	K_1J_1	K_2J_1	K_3J_1	K_4J_1	K_5J_1	K_6J_1	K_7J_1	K_8J_1
Social Dimension	K_1J_2	K_2J_2	K_3J_2	K_4J_2	K_5J_2	K_6J_2	K_7J_2	K_8J_2
Economic Dimension	K_1J_3	K_2J_3	K_3J_3	K_4J_3	K_5J_3	K_6J_3	K_7J_3	K_8J_3

Table 3-1: Multi-Dimensional Representation of the Cyber-Attack Kill-Chain (Source – Author)

Or

$$\begin{bmatrix} K_1J_1 & K_2J_1 & K_3J_1 & K_4J_1 & K_5J_1 & K_6J_1 & K_7J_1 & K_8J_1 \\ K_1J_2 & K_2J_2 & K_3J_2 & K_4J_2 & K_5J_2 & K_6J_2 & K_7J_2 & K_8J_2 \\ K_1J_3 & K_2J_3 & K_3J_3 & K_4J_3 & K_5J_3 & K_6J_3 & K_7J_3 & K_8J_3 \end{bmatrix}$$

Where K_1J_1 is a structural representation of multi-variate time series vectors, predictive of the antecedents to said cyber-attack on the physical dimension.

This research puts forward that events at each stage of the kill-chain can be characterized across the information space. The cross-dimensional matrix shown in table 3-1 above is a representation of the various possible dimension-phase combinations of the theorized approach. The matrix is a decomposition of the various phases of a hypothetical kill-chain. Each kill-chain must be put in context and sources of evidence identified accordingly across the information. Every cell in the corresponding matrix is a possible path identifying the evidence sources for events that characterize that phase of the kill-chain in a given cyber-attack scenario. This matrix is useful in tracking-in the context of a cyber-attack-where in the kill-chain an adversary is and what dimension of the information space characterises the adversaries activities at that phase.

3.3.2 Hypothesis Two

Given a set of multivariate time series, observations of a set of variables representative of events on the physical, social or economic dimension of cyberspace, Xt, Yt and Zt, the intra-dimensional co-integrating rank between Xt and/or Yt and/or Zt is greater than or equal to 1. i.e. r >= 1.

The resulting cross-dimensional multi-phased matrix representation is further analysed using a structural VAR/VECM model. The VAR/VECM model at each stage of the kill-chain is constructed using features that are predictive of activities that characterise the current stage of the kill-chain with early indicators for the next phase of the kill-chain. Incorporating the kill-chain produces a VAR/VECM representation of the physical-social-economic cyber kill-chain that is capable of filtering down to multi-variate predictive features represented in a structural VAR/VECM model.

3.3.3 Hypothesis Three

The events of the cyber-kill chain which represent activities of the attacker, can be attributed to eitherX_t, Y_t or Z_t. i.e we can formulate a multi-dimensional physical-social-economic kill chain using events characterized as occurring on X_t, Y_t or Z_t.

3.4 ENHANCED FRAMEWORK FOR THE ENTANGLED CYBERSPACE

As earlier stated, let $\boldsymbol{X_t}$ represent information on the physical dimension of cyberspace, $\boldsymbol{Y_t}$ represent information on the social dimension of cyberspace and Zt represent information on the economic dimension of cyberspace, at each stage of the kill-chain, data is pulled from respective dimensions to construct a VAR(p) model at time t to t-n. As theorized above, assuming events on the cyber-kill chain occur across various dimensions of cyberspace, we derive subsets of $\boldsymbol{X_t, Y_t}$ $\boldsymbol{and}$ $\boldsymbol{Z_t}$ across cyberspace. Following the model presented in **Error! Reference source not found.** above, we create the following structural model for a VAR/VECM representation of the cross-dimensional multi-phased framework with a 9-phased cyber-kill chain as shown in section 2.8.2. The phases in any specifc kill-chain correspond with the number of individual activities an attacker needs to take to implement that attack. Therefore, the number of phases for various attack kill-chains may vary slightly.

The enhanced entangled cyberspace theory is built taking the logical sequence of a cyber-attack perpetration into consideration. It implements the traditional kill-chain across the various dimensions of cyberspace by constructing a structural predictive model at each phase of the kill-chain. Therefore, each phase of the kill-chain is attributed to a single dimension as a major source of evidence.

3.5 CONCLUSION

This chapter has discussed the development and mapping of the entangled cyberspace framework in alignment with theory. It further discusses the techniques for practical implementation of the vector autoregressive models for simultaneously investigating multiple time-based evidences from multiple dimensions of cyberspace. This chapter provides a mapping from a summarized view of the entangled cyberspace model; as it may be useful to strategic decision makers in cyber security to an in-depth view of implementation as will be useful to technical analysts. In summary, this chapter deconstructs the process of untangling the complex dynamics of events in given webbed data.

4 CHAPTER 4: RESEARCH METHODOLOGY

4.1 INTRODUCTION

This chapter provides validation to the research approach followed in investigating the behavioural approach through which cyber-incidents can be predicted in cyberspace. This chapter contains detailed description of the research methodology, statistical methods and technical solutions used in developing the conceptual framework, testing the theoretical models and implementing the research artefact. The theory of the entangled cyberspace is built on a structural, behavioural predictive model and demonstrated by testing and evaluating the models on real-world datasets simulated to contain the theorised entanglements. In this research, the general methodological approach is the development of an experimental scenario to test the concepts of the developed framework. This research builds an experimental environment with simulated datasets provided to capture the nature of an entangled cyberspace. This research works under the assumption that the developed model can pick up the nature of entanglements built into the experiments. The experiments and scenarios are set up using benchmark datasets simulated outside the framework of this research but containing the intended entanglements of interest. Using an externally developed benchmark dataset to test the conceptual model ensures the absence of bias as datasets have been developed outside the research point of view and therefore not forced to contain relationships of interest.

This research employs a behavioural research paradigm and a design science research paradigm. The behavioural research paradigm focuses on understanding the behaviour of actors in the entangled cyberspace. Actors such as the attackers, the targets and the analysts portray certain behaviour in various scenarios that affect the nature of inter-relationships in the entangled cyberspace. Any solution involving these actors must also address frameworks for describing, explaining and therefore predicting behaviours of these actors. Incorporating a behavioural paradigm into the research environment ensures that behavioural elements that act as evidence for establishing relationships between entities in cyberspace are appropriately captured by our model. The design science paradigm ensures that all research artefacts are useful in addressing the intended objectives of the research. This research is based on a mixed methods approach that ties both paradigms together using both qualitative and quantitative analytical approaches to test and evaluate the theory put forward.

4.2 OBJECTIVES AND OVERVIEW

This chapter provides a critical review of the plausible research methodologies that apply to the objectives put forward in this research. The research approach used in this thesis is a combination of methods to achieve the research aims and objectives. The aim stems from the overall goal of building predictive capacity that contributes to the area of cyber situational awareness, incident readiness and proactive cyber defence.

4.2.1 Aim

Develop an approach and theoretical framework for detecting early warning signals of cyber-attacks and therefore predicting cyber-attack attacks in the complex cyberspace. This is achieved by characterising and extracting indicators of cyber-attacks across the dimensions of cyberspace to accurately feed statistical and mathematical models for prediction. The theoretical aim of this research involves conceptualising the nature of entanglements in cyberspace. The technical aim of this research

involves building a complex structural model to handle the challenge of untangling the entanglements in cyberspace. Furthermore, the researcher hopes to develop a research artefact to demonstrate the theories in the conceptual framework.

4.2.2 Objectives

a) *To characterise cyberspace as a collection of multiple dimensions that are part of a whole.*

b) *To determine the features that define the nature of entanglements in cyberspace.*

c) *Propose a multi-dimensional structural framework for integrating multiple sources of evidence in cyberspace to pre-empt cyber-attacks.*

d) *Evaluate the methods and algorithms for an integrated approach to pre-empting cyber-attacks.*

e) *Assess the implications of the research findings and an integrated approach for pre-empting cyber-attacks in practice.*

This chapter discusses in detail the methods employed and provides a suitable justification for their use. The methods employed for addressing each of the stated objectives are broken down and critically discussed.

4.2.3 Methodological Review

This project aims to develop a comprehensive, intelligence-driven and accurate approach for pre-empting cyber-attacks with the goal of achieving intelligent cyber situational awareness and proactive cyber defence. In response to the research demand, the researcher analysed various research methodologies, eventually opting for a mixed method approach. The mixed method approach enables a combination of quantitative and qualitative methods; thus, its efficiency has been proven in social sciences (Creswell, 2003). The mixed method approach was selected for several reasons.

The quantitative methods seek to employ deductive reasoning by understanding the causes of a phenomenon thus attempts to verify its effects on its surrounding environment through mathematical and statistically valid methods (Creswell, 2003). In the context of this thesis, quantitative methods are important drivers for this research as they provide a method for evaluating said dependences in cyberspace (Hevner *et al.*, 2004). Quantitative methods also form the basis for the identification of entanglements in cyberspace, i.e. to evaluate the cyber-attack predictive capacity of the entangled cyberspace and its ability to increase cyber situational agility.

The combination of both methods (qualitative and quantitative methods) offer opportunities for exploiting the design science approach offered by (Barnett, Smith and Whittington, 2014). This approach will also ensure that the researcher's philosophical standpoint is represented in the research and backed-up with concrete proof. It ensures that the technical artefact is based on philosophical aspects of real-world problems and is geared towards addressing these problems.

Also, included in the research methodology is the epistemology; the body of knowledge the researcher includes in the theory. Such theories of knowledge include Objectivism or Subjectivism which inform the researcher's philosophical standpoint on some philosophical research systems such as positivism, interpretivist or critical theory (Creswell, 2003). (Creswell, 2003) describes interrelated levels of

decisions based on the above perspectives, which sets up the steps for the research design with three basic questions:

a. What are knowledge claims being made by the researcher especially from a theoretical perspective?
b. What strategies of inquiry will inform the procedures and techniques used in the research?
c. What methods will be employed for data collection, integration and analysis?

4.3 RESEARCH PARADIGMS FOR IS RESEARCH

One common logic used to streamline Information Systems research paradigms is the differentiation between behavioural science and design science (Davis, 2000). While design science research seeks to produce technological artefacts to solve organisational problems, behavioural sciences focus on the development and justification of theories to understand why things are the way they are. Behavioural science is therefore referred to as the 'problem understanding paradigm' while design science is referred to as the 'problem-solving paradigm'.

	Behavioral Sciences	Design Sciences
Origin	Natural Science	Artificial Science
Paradigm	Problem Understanding Paradigm	Problem Solving Paradigm
Objective	Develop and justify theories that explain or predict human phenomena surrounding the analysis, design, implementation, management, and use of information systems	Create innovations that define ideas, practices, technical capabilities, and product through the analysis, design, implementation, management, and use of information systems.
Object	Human-Computer-Interaction	IT Artefact Design

Table 4-1: Design Science and Behavioural Science

4.3.1 Behavioural Science Research Framework

Behavioural research focuses primarily on understanding patterns of behaviour in a given system. According to (March and Smith, 1995), behavioural sciences is one of the many domains of natural sciences. Natural sciences include traditional research into physical, biological, social and behavioural domains aimed at understanding reality. According to (Leary, 2012), the behavioural scientist's job is to detect and explain phenomena. Natural scientists develop sets of concepts or specialised language, with which to characterise phenomena. These developed concepts are further used in higher order constructions such as laws, models and theories, to make claims about the real nature of reality. Theories – deep, principled explanations of phenomena, are the crowning achievements of natural and behavioural science research. Products of behavioural sciences are evaluated against the norms of truth or explanatory power. Claims must be consistent with observed facts, the ability to predict observations based on a study of past behaviour is an example of explanatory success. (Leary, 2012) identifies three main goals of conducting behavioural science research.

4.3.1.1 Describing behaviour

Behavioural research sometimes focuses primarily on describing patterns of behaviour, thought, or emotion. For example, natural science researchers conduct large studies of randomly selected respondents to determine what people think, feel and do. Public opinion polls, such as those that dominate the news during elections, that describe people's attitudes and preferences for candidates. Research in clinical psychology and psychiatry investigates the prevalence of certain psychological disorders. In the field of marketing and sales, researchers study consumers' preferences and buying practices.

4.3.1.2 Predicting Behaviour

Researchers are sometimes interested in predicting people's behaviour or the behaviour of a system after a certain period. Personnel psychologists try to predict employees' job performance from employment tests and interviews. Similarly, psychologists attempt to predict academic performance from scores on standardised tests to identify students who might have learning difficulties in school. Likewise, some criminal forensic psychologists are interested in understanding variables that predict which criminals are likely to be dangerous if released from prison. Economists are also interested in predicting the behaviour of the macroeconomic indicators at some point in the future to guide policy decision. Financial analysts also predict movements of financial time series to guide investment portfolios. The tests to be used must be administered, analysed, and refined to meet certain statistical criteria. Then data are collected and analysed to identify the best predictors of the target behaviour. Prediction equations are calculated and validated on other samples of participants to verify that they predict the behaviour successfully. All along the way, the scientific prediction of behaviour involves behavioural research methods.

4.3.1.3 Explaining Behaviour

Some researchers in the field of behavioural sciences regard 'explanation' as the most important goal of scientific research. Although description and prediction are quite important, scientists usually do not feel that they understand something until they can explain it. We may be able to describe patterns of violence among prisoners who are released from prison and even identify variables that allow us to predict, within limits, which prisoners are likely to be violent once released. However, until we can explain why certain prisoners are violent, and others are not, the picture is not complete.
Furthermore, research in natural sciences descriptive, correlational or experimental. Descriptive research seeks to describe thoughts, feelings, and trends and seeks to answer the question of why systems, people or individuals act the way they do. Correlational research seeks to investigate relationships and dependence among a certain set of identified features in a system. The experimental research investigates cause and effects of changes in patterns, behaviours or trends.

This research combines these two behavioural frameworks discussed above to produce a holistic approach to addressing behavioural effects in the research experiment. The rationale behind combining these frameworks are as follows:

i. To properly conceptualise the behaviour of interest as relevant to the research aims and objectives.
ii. To properly identify all actors involved in a specific behaviour of interest.

iii. To properly capture the dynamics of the domain in which the actors act and the observed behaviour occurs.
iv. To ensure that experiments are geared towards understanding, explaining and describing all external and internal factors that enact, enable, cause and affect the behaviours of interest.
v. To properly identify all behaviours and actors linked to behaviours of interest under observation.

	Identify	Describe	Predict	Explain
Behaviour	A Cyber-Attack	The Cyber-Kill Chain	The type of Cyber-attack	
Actor	Attacker	Individual, Hacker groups, State Actors		
Domain	Cyberspace	A multi-dimensional space comprising of individual layers acting independently and as a whole.		
Durability	One-off, Repeated, Dependent or Enduring			
Scope	Inter-related			

Table 4-2: Combined behavioural framework

The table above characterises the concepts built into the experiment.

4.3.2 Design Science Research Framework

(Hevner *et al.*, 2004) design science approach for information systems research has produced a seven-point guideline in which they write that knowledge and understanding of a design problem and its solutions are acquired in the building and application of an artefact, as shown in the table below:

Design Science Guideline	Description	Research Alignment
Design as an Artefact	Design science research use produce a viable artefact in the form of a construct, a model a method or an instantiation.	This research presents a framework for a holistic approach to pre-empting cyber-attacks in a multi-dimensional cyberspace. Furthermore, this research produces a technical artefact to demonstrate the theories put forward.
Problem	The objective of the design science research is to	This research addresses a

Relevance	develop technology-based solutions to important and relevant business problems in the real-world.	fraction of the problem of proactive cyber situational awareness by providing a framework that puts defenders ahead.
Design Evaluation	The utility, quality and efficacy of a design artefact must be rigorously demonstrated with a well-executed evaluation method. The evaluation methods proposed can be further broken down into the following groups: • Observational: using case studies in business environments or field studies. • Analytical: Examining the artefact for static qualities such as complexity, fit of artefact, or its performance. • Experimental: Using controlled experiments or simulation where one can run the artefact with simulated data. • Testing: Black Box or White Box testing. • Descriptive: Use of information from knowledge base to build argument for artefact's utility or in the form of scenarios.	This framework is tested using the well-established cyber kill-chain, by predicting each phase on the kill chain. Each component in the framework is tested real-life experimental scenarios using benchmark datasets.
Research Contributions	Effective design science research must provide clear and verifiable contributions in the areas of the design artefact, design foundations and/or design methodologies. • Design Artefact –The ability to reuse the artefact itself to solve other unsolved problems. • Foundations: To create new constructs, models, methods etc. to also extend or improve existing foundations. • Methodologies: to produce new ways to evaluate and create new contributions to design science. An example is a framework for predicting and explaining why a particular information system will or will not be accepted in a given organisational setting (Venkatesh, 2000).	The project would contribute specifically to the area of cyber situational awareness and predictive analytics in cyber defence. It combines various established tools and techniques to create a unified process for pre-empting cyber-attacks in cyberspace.
Research Rigor	Design Science research relies upon the application of rigorous methods in both the construction and evaluation of the design artefact.	The literature review is based on verified academic and professional sources. The methods, techniques and data

		used in this research have been proven and tested in academia.
Communication Of Research	Design science research must be presented effectively both to the technology-oriented as well as the management-oriented audiences.	The research consists of Data summary facts and figures based on analysis of datasheets and variables explained in the data analysis.

Table 4-3: Design Science Research Framework

Concerning the points outlined above, (Hevner *et al.*, 2004) explains that, although the creation of an innovative, purposeful artefact for a specific domain is applied usefully, the artefact must also yield utility for the specified problem. Therefore, a thorough evaluation of the artefact is crucial. Novelty is similarly crucial, since the artefact must be innovative, solving an unsolved problem or solving a known problem more efficiently and/or effectively. This concept differentiates design science paradigm from the practice of design. The artefact must be rigorously defined, formally presented, coherently and internally consistent. The process by which the artefact is created uses a search process whereby a problem space is constructed, and a mechanism proposed or enacted to find an effective solution. Finally, the results of the design science research process must be communicated effectively to all kinds of audience interacting directly with the artefact.

4.3.3 Combining Behavioural and Design Science

In an attempt to ensure that IS research is thoroughly constructed to meet human needs, March and Smith also recommended the combination of natural/behavioural science and design science. Researchers often face the challenge of conceptualising, representing and selecting appropriate techniques for solving real-world problems. If IS research is to make significant progress IS research methods must include in its process, mechanisms for developing an understanding of how and why IS systems do or do not work for target populations. Therefore, the following framework for undertaking IS research that combines both sciences is proposed.

	Build	**Evaluate**	**Theorise**	**Justify**
Constructs: **conceptualisation of a domain space used to describe the problems that exist within that domain and specify their solutions**				
Models: **a set of propositions or statements expressing relationships among various specified constructs.**				
Methods: **a set of steps (algorithm or guideline) used to perform a given task.**				
Instantiation: **the realisation of an artefact in the problem domain.**				

Table 4-4: IS Research for Behavioural and Design Science

It is important to state that behavioural science is at the core of developing the theory of complexity in cyberspace, while design science develops the lab experiments, test beds and research artefact to put the theory of the Entangled cyberspace into practice. Consequently, combining both research paradigms are based on evidence form (March and Smith, 1995), who pointed out the importance of collaboration between both sciences. (March and Smith, 1995) propose that:

- Design science creates artefacts, giving rise to phenomena that can be the targets of behavioural science research. Group decision support systems foster user behaviours that are the subject of behavioural science investigations.
- Since artefacts have no dispensation to ignore or violate natural laws, their design can be aided by an explicit understanding of natural phenomena. Thus, natural scientists create knowledge which design scientists can exploit in their attempts to develop technology that meet human needs. In conclusion, (March and Smith, 1995) provide an example with a healthcare example pointing out that the explanation of why a drug is effective in combating a disease may not be known until a considerable time lag after the drug is in common use.

4.4 SOCIAL-POLITICAL-ECONOMIC-CULTURAL (SPEC) EVENT ANALYSIS

The most significant component of cyberspace are the humans involved. Current cyber prediction models are restricted to analysis on the network layer of cyberspace, but as shown in (Gandhi *et al.*, 2011a), these methods fail to account for the impact of human behaviour in the process of a cyber-attack. Evidence (Gandhi *et al.*, 2011a; Bronk and Tikk-Ringas, 2013; Hernández *et al.*, 2016) has shown that more and more cyber-attacks in recent times are linked to social, political, economic and cultural events. The adversary's level of sophistication and their motivations are essential to predicting, preventing, tracing and attributing cyber-attacks. As a result, SPEC factors have the potential to be early predictors of surety events in cyberspace.

Consequently, methods for analysing correlations between these types of events in cyberspace have become useful for providing valuable insights regarding agents, motivation and means of cyber-attacks. For example, (Bollen and Mao, 2011) found a connection between social media discussion and fluctuations observed in the stock market. Hernández et al. (Hernández *et al.*, 2016) also predicted cyber incidents by analysing twitter discussions. Lastly, (Chakraborty *et al.*, 2016) was able to predict socio-economic events by analysing news articles and extracting events. SPEC events present an analytical taxonomy along SPEC interdependence dimensions.

4.5 SCENARIO DEVELOPMENT

This section explains the process used in developing the scenario on which the experiment in this research is based. Maier et al. (Maier *et al.*, 2016) highlight the importance of uncertainty in the future due to rapid changes in the physical reality which has led to the need for methods which requires uncertainties to be described with the aims of scenarios that represent coherent future pathways based on a different set of pre-assumptions. Eriksson et al. (Eriksson, Olofsson and Ekvall, 2003) describe a scenario as a picture of the future conditions of an object within its environments. The 'conditions' in this instance refers to characteristics of the results of a given sequence of events (situations) and factors which disturb the natural evolution of these occurrences. (Bishop, Hines and Collins, 2007) describes a scenario as a description of plausible situations and what they might lead to. Generally, (Börjeson *et al.*, 2006) presents three main categories of scenario studies based on questions a user may want to pose about the future.

Explorative scenarios answer the question 'what can happen?'. The aim of designing explorative scenarios is to explore situations that are regarded as possible to happen from a variety of perspectives. Typically, a set of scenarios are developed to span a wide scope of developments. Explorative scenarios are used for long-term observations and allow for structural and profound changes in the modelled system.

Normative scenarios answer the question of how a specific target can be reached based on how the system is treated. The aim of designing normative scenarios is to test how a certain target can be reached by adjusting the current situation or by changing the structural blocks on which the system depends.

Predictive scenarios present a detailed and quantitative indication of how the system will change under a certain set of conditions. This type of scenario development experiments usually answers the question 'what will happen?'. These types of scenario development are done in an attempt to predict the future. Predictive scenarios are usually developed under the assumption that the laws governing a given predicted system would remain constant and unchanged over the predictable time. This scenario is explored as part of this research as it concerns the forecasting of the future behaviour of a system under certain conditions.

Additionally, the (Maier *et al.*, 2016) puts forward a set of processes for developing a scenario. These steps are illustrated in the table below.

S/N	Scenario Development Step	Description	Research Application
1	Scenario Domain Definition	This defines the domain around which the scenario is built.	The scenario covers events in cyberspace as analysed from its three identified dimensions Physical, Social and Economic.
2	Description of Key Events	This outline the key events that link the scenario to theory.	Three events were designed to model entanglements in cyberspace using simulated data.
3	Description of Key Measurements	This identifies the key measurements for the justification of the proposed theory.	The key measurements for the confirmation of theory include co-integration across the layers of cyberspace and a measurement of causality.

Table 4-5: Framework For Scenario Development. Source (Maier *et al.*, 2016)

The aims of the scenario development process are as follows:

a. To model representative data from the multiple dimensions of cyberspace.
b. To simulate inter-dependent relationships between representative data across the multiple dimensions of cyberspace.
c. To test for co-integrating and causal relationships between representative data across the multiple dimensions of cyberspace.
d. To predict the occurrence of a cyber-attack on the physical dimension.
e. To develop a generalised form of the Entangled Cyberspace using a predictive model for APT cyber-attack using features from the multi-dimensional Cyberspace.

4.5.1 Validating the Scenario

The proposed scenario would need to address each of the '5Cs' listed above to ensure that scenario environment is representative of real-world operating environments across all domains. This will, in turn, lead to a validated approach. The key goals of the scenario are:

- To simulate inter-dependent relationships between representative data across the multiple dimensions of cyberspace.
- To extract some features across all dimensions of cyberspace that are predictive of a cyber-attack on the network layer.
- To predict with certain precision, the occurrence of a cyber-attack in the Network Layer of cyberspace.

The mission of this experiment is to predict the occurrence of a cyber-attack on the network layer of cyberspace. It should also be able to handle the different formats of data and integration requirements needed. The objective of developing this scenario is to create a representative environment for an Entangled cyberspace to test new predictive models for cyber-attacks in cyberspace. Therefore, this scenario is designed with some basic assumptions. Firstly, the scenario is designed based on the pre-assumption of inter-dependence between features in the scenario design. Secondly, the scenario assumes a sequential course of events leading up to a cyber-attack on the network layer of cyberspace.

4.5.2 The Scenario

The scenario used in this thesis is based on the Vast IEEE 2011 mini-challenge one and mini-challenge two (Cook *et al.*, 2011). In this thesis, the researcher uses a mirrored scenario designed around a fictional company, Delish Corporation in the fictional city of Zuma. Delish Corporation was founded by Mr Darren Peters and has operated in the Zuma area for eight years since it started operations in 2009. Delish Corporation is an agricultural company dedicated to providing consumer products across its various outlets. Over the past five years, Delish has recorded huge amounts of profits with their stock market prices hitting an all-time high in 2013. The scenario is further broken down into three sub-events surrounding the subject: An epidemic spread in the city of Zuma, a stock market crisis relating to Delish Corporation stocks and a data exfiltration attack on Delish Corporation. Each event in the selected scenario are linked to each other. The scenario data is designed (Cook *et al.*, 2011), to contain information about linked activities. Although previous legacy datasets for building cyber defence models such as DARPA (Mahoney, 2003b; Thomas, Sharma and Balakrishnan, 2008), KDDCUP (Gehrke, Ginsparg and Kleinberg, 2007; Zhang and Wang, 2013) and VERIS (Creswell, 2003; Liu *et al.*, 2015), activities within these datasets are mono-dimensional, limited to the network layer of the information space. This data is different from other datasets prior presented in that it cuts across the information space. The dataset provides a representation of theorized entanglements which makes it suitable for the experimental design.

4.5.2.1 Event 1: The Epidemic Spread

Zuma is an urban city with a population of 9 million people. During the last few days, health professionals at local hospitals have noticed a dramatic increase in reported illnesses. Observed symptoms are flulike and include fever, chills, sweats, aches and pains, fatigue, coughing, breathing difficulty, nausea and vomiting, diarrhoea, and enlarged lymph nodes. More recently, there have been several deaths believed to be associated with the current outbreak. The problem made national headlines when a former military veteran dies from these poisoning symptoms. His death sparks an unexpected reaction in the populace who are demanding an answer from state officials.

4.5.2.2 Event 2-Delish Stock Prices Crash

Delish Corporation is a Limited Liability company listed on the stock market. Delish has enjoyed souring prices and maximum returns over the last three years. However, in the last eight months, there has been a gradual decrease in the recorded market capitalisation index of the company's shares. Additionally, there was a major incident in the officially recorded stock market prices sometime between February 2017 and May 2017. The incident sent the prices of Delish corporation shares slumping suddenly. The sudden change in the value of the company's shares sets of a continued decreasing trend in the Company's share prices over the next five months. Financial officials have attributed this fall in prices to the negative press, reduced client loyalty, reduced client trust and loss of investor confidence in Delish Corporation.

4.5.2.3 Event 3- Cyber Attack On Delish Corporation

To support its production, marketing, sales, logistics and distribution, Delish Corporation operates a fully integrated IT infrastructure. The company operates a central server for storing data such as transaction figures (both online and in-store), customer account details, credit/debit card details, emails etc. The company's IT infrastructure is managed by an in-house cyber-savvy team, responsible for Monitoring, Detecting and Preventing any perceived cyber incident of the company's servers or any subnetwork. The team is also responsible for responding and mitigating the risks of cyber incidents in the event of their occurrence.
Critical assets relating to product patents and trade secrets are stored on specific hosts within the organisation's network. Lately, an IT staff has complained about the increasing rate of anomalies in the company's weblog and network traffic. Due to the critical timing of these inconsistencies, this was brought to the attention of the company's CEO, Mr Darren Peters. Mr Darren suspects a connection of this incident to the recent crash in the company's market shares and orders a re-evaluation of the company-wide network.

The activities in the scenario presented represents a series of linked events as described in sections 4.5.3.1-4.5.3.3. The scenario is based in the following assumptions of connections between these events:

- Delish Corporation is the major player of all agricultural and eco-related activities within the metropolitan city of Zuma.
- The operational activities of Delish Corporation have slowly over the preceding years led to land, air and water poisoning in the metropolitan area.
- Delish Corporation operates a fully integrated IT network infrastructure that supports its business operations.
- There has been previous complains and legal actions against the activities of Delish within the metropolitan area.
- This has led to a negative public perception of the company that affects the values of their stock prices.
- Disgruntled members of the community launch a cyber-attack on delish corporation networks mainly to disrupt its activities and make a political and economic statement.

4.6 DATA GATHERING TECHNIQUES ADOPTED IN THIS THESIS

This section discusses the techniques used for collecting data used in this research. A critical process for any research project is data identification and collection through inquiry methods that are guided by the purpose of the research and which are influenced by the investigation of the researcher. The data collection method used in this research is accordingly chosen based on the research objectives. This research uses data collected from scenario simulations, lab experiments and data simulations. Lab experiments were created for each event (event 1, event 2, and event 3) in the scenario. Data for event 1 and event 2 of the scenario was generated by scientifically simulating co-integrated time series (Galenko *et al.*, 2009). Social media data for event two was based on a benchmark data set in (Schreck and Keim, 2013). Timestamps that fell outside the timeframe of the scenario analysis, were ignored for this experiment. Data for event three was generated in a cyber range that simulates the network infrastructure defined in the scenario. The cyber range was designed and developed so the inter-dependence between features in cyberspace can be identified and validated. A holistic experimental process that ties the three events together is followed by application of quantitative methods for testing inter-dependence in time series. This section further expatiates on the data gathering, collection and simulation techniques used in this thesis.

4.6.1 Using Simulated data for research

One of the major challenges in analytical research is the acquisition of large representative data sets that model a given hypothetical theory. Simulations, as defined by the Webster Dictionary refers to an imitation, well-constructed, to pass for the real thing. Data simulation refers to the imitation of relationships amongst features to create datasets that follow a particular distribution and are representative of behaviours researchers hope to test (Thesen and Travis, 1990). Data simulation provides more robust solutions by offering a certain level of knowledge and control in the experimentation process. In research design, (Thesen and Travis, 1990) conclude that simulations are useful for:

- Serving as a tool to help address the complex interaction of data construction, analysis and statistical theory.
- Improving researchers' understanding of basic research principles and analytical techniques.
- Investigating the effects of a given behaviour on another behaviour.
- Exploring the accuracy of novel analytical techniques applied to data structures.

In this research, the researcher first creates data according to given theoretical models and further examines how well the concepts can be detected using the techniques proposed in this research.
Building good simulated data is the most critical step and is often challenging. As (Sharma *et al.*, 2010) point out, simulated data which are too clean or too fitted will provide misleading or inconclusive results. On the other hand, simulated data with large errors and numbers of outliers, which do an not correctly model the behaviour of interest, will also provide misleading or wrong results. Benchmarking is proposed as a solution to generating realistic datasets from simulation efforts. The process of benchmarking compares the simulated dataset to real-world problems that are widely accepted and understood to improve the data's credibility. This may be done by considering historic events in the field of research, applications of such techniques and specific industry under consideration. Additionally, precedent examples in the bench, mark or historically significant cases that are characteristic of the problem under investigation may also be considered.

4.6.2 Lab Experiments

Laboratory experiments are controlled activities that create and observe the behaviour or event of interest. The logic of laboratory experimentation is that it is controlled therefore enabling researchers to precisely measure the effects of exogenous variables on endogenous variables, thus establishing cause and effect relationships. Consequently, predictions about future behaviour of variables in the system can be made under certain assumptions. The lack of data for cyber defence research is a looming problem identified by multiple researchers. This stems from the industry's unwillingness to share information related to cyber incidents. Factors for this mistrust have been identified as protection of business integrity, protection of business revenue model, protection against cooperate espionage and corporate sabotage amongst many others. This has led to the need for an elaborate solution to understanding cyber-physical networks. The logic behind Laboratory experiments in the area of cyber defence is to re-create cyber-physical networks in a controlled environment and re-construct network behaviour under certain circumstances. For example, observing the pattern of network traffic in the event of a request overload. Controlled environments created for cyber defence research are called 'Cyber Ranges' or 'Cyber Testbeds' and may be physical or virtual. This research uses data collected in a physical, cyber range, constructed to reproduce a distributed denial of service attacks and port scan attacks as outlined in the scenario design.

4.7 EXPERIMENTAL DESIGN

This section outlines the process adopted in designing the experiment. This section outlines the process adopted in designing the experiment to prove the concepts presented in this research. The experiment is designed to address the identified research problem of detecting early warning signs of cyber-attacks and therefore pre-empting cyber incidents in cyberspace. The experimental design begins with an outline of its intended aims and objectives. It further presents the development of the experimental scenarios and the experimental sequence. Additionally, this section critically evaluates the experimental testing environments and its alignment to the overall research aims and objectives.

4.7.1 Experiment Objectives

This experiment aims to test the predictability of cyber-incidents in cyberspace using a multi-dimensional structural integration approach by fusing data streams across multiple dimensions of cyberspace. The experiments objectives are designed to address the overall research aim. The main objectives are as follows:

1. Identify indicators of each phase of the cyber-kill chain on the multi-dimensional cyberspace with the aim of pre-empting the last phase of the kill-chain.
2. Characterise the Cyber-Physical-Social-Economic Kill-Chain in the multi-dimensional cyberspace using structural statistical techniques.
3. Identify the variables that actively contribute to pre-empting the phases of a cyber kill-chain during the perpetration of a cyber-attack.
4. Test the efficiency of the proposed entangled cyberspace model in pre-empting cyber-attacks.

4.7.2 Design and Development of Research testing environment

This part of this section covers all aspect of developing the testing environment that meets the research requirements. The social-physical-economic cyber kill-chain, the multi-dimensional cyberspace and Vector Auto-Regressive Models informed the development of this research's testing environment directly by providing the main factors and components required for constructing the

entangled cyberspace models. Additionally, the events described in the scenario section 4.5.2 of this research inform the experimental development by identifying the features and capabilities to be included in the testing environment. This section will provide more insight into the testing environment, its components and infrastructure.

Systems Requirements: The required system can be described as an advanced intelligent predictive system, using multiple sources of evidence to deliver high-quality information in an inter-connected space. This system which has the potential to improve cyber situational awareness to cyber-attack defenders has three main components:

- The multi-dimensional cyberspace which is further divided into the physical, social (Barnett, Smith and Whittington, 2014) and economic (Gandhi *et al.*, 2011a) dimensions with multiple sub-layers each as explained in chapter 2 of this thesis.
- The cyber-attack kill chain which models the steps an adversary must take achieve a successful attack on a target network. The cyber-attack modelled in this research is a Data Exfiltration attack as explained in chapter 3 of this thesis.
- Statistical predictive techniques. The default structural analytical tool used is the Vector Autoregressive Models (Luetkepohl, 2011). However, the model is robust to accommodate optimal solutions at each phase of the kill-chain. Therefore, additional structural techniques such as Vector Error Correction Models (Sims, 1980) are also considered in the model.

The system provides predictive solutions that captures the subject's cyber operating environment in the context of the entangled cyberspace while optimising the nature of available evidence and the observed nature of entanglements. The system is also designed to enhance cyber situational awareness level of the cyber defender from a multi-dimensional perspective. It is hypothesized that active components in the system come together to provide a predictive capacity that keeps defenders ahead of the threat using evidence from multiple domains.

4.7.2.1 Requirements For Testing Environment

Based on the discussion provided so far in this chapter, to test the Entangled Cyberspace Theory, the following requirements should be considered while developing the testing environment:

1. A computer network infrastructure that reflects the physical assets described in the scenario.
2. A chosen timeframe for analysis of scenario events.
3. Simulated network and system control capabilities.
4. Simulated network data, e.g. network flow, ids logs and security logs - with real activity within the chosen time frame collected from the computer network infrastructure.
5. Simulated social forum activity within the chosen time frame.
6. A simulated economic, political or cultural activity within the chosen time frame
7. Identification of the simulated phases of the cyber-attack kill-chain across the dimensions of cyberspace.

To achieve the above stated, the system requires highly sophisticated technologies and analytical capabilities that can provide real-time analysis of evidence sources in the context of the cyber-attack. This is achieved by integrating different types of technologies and tools that offer real-time analytical capabilities. The use of a predictive analytical environment is crucial for cyber defenders as such a platform would allow evidence feeds to be integrated across multiple domains on a time-dependent scale. Due to these requirements, the researcher includes the following components in the system:

- A social micro-blogging platform with active participants: This platform is set up to simulate discussions amongst participants that should be predictive of cyber-incidents. These

discussions are set to take place on the social dimension of cyberspace capturing the thoughts of participants on a specific topic. Simulating the dynamics of cyber-related discussions is paramount, especially within the given time frame.

- A target network: The simulation of a target network gives the researcher the ability to act in a real-world look alike network with similar network activities occurring. Such network activities provide the researcher with knowledge of signatures that characterise the different phases of the kill-chain on the network layer of cyberspace. The network simulates a target's network and the infrastructure and services running within it.
- A financial crisis: The simulation of a financial crisis about a specific subject of analysis gives the researcher the ability to generate synthetic financial data using a combination of tools. The first necessity is a sample of real financial stock market prices as a training model from which our synthetic financial data would be generated. The main objective of including financial data is to represent activities on the economic dimension of cyberspace.
- An Attacker: The attacker in this experiment is the enemy. The enemy may be an individual or group of individuals. The main objective of including the attacker in the experiment is to penetrate the testing environment while leaving behind subtle indicators for identification by research tools. The expectation lies in the ability of the theoretical model to identify these indicators in the test environment.

4.7.3 Operations of Testing Environment

The operating environment within which the experiment is designed is explained in this sub-section. The environment is designed to simulate all inter-connected events within the stipulated time frame analysis. The environment models three major events: epidemic spread in the fictitious city of Zuma, Delish Corporation stock price crisis and a data exfiltration attack on Delish Corporation.

4.7.3.1 Evidence Identification and Monitoring On the Social Dimension (Epidemic Spread)

The system operates under the assumption of an embedded social platform for capturing discussions in social networks related to a subject interest. Cyber analytics teams are responsible for identifying such communication channels and providing a clear access to information generated on these platforms. Cyber defence strategies should utilise all skills, tools and technologies to gather intelligence from all enemy domains and provide real-time actionable intelligence. Also, cyber analysts in the system are in charge of implementing a security management framework that includes proactive techniques such as prevention, monitoring, analysis and detection to stay ahead of adversary tactics. On the other hand, there is also a responsibility on the part of cyber operators to identify and monitor all end-users of a cyber network in social forums and communication channels. Such endeavours require a certain degree of social forum analysis techniques such as proposed by (Matusitz, 2011; Schreck and Keim, 2013). The available communication channels in the simulated system are assumed to be sufficient in capturing the required inter-relationships between entities in the system. Also, the communication feeds used in this research is simulated to provide the necessary intelligence that covers activities on the social, physical and economic dimensions of cyberspace in the context of the scenario.

Additionally, communication feeds on the social dimension have been proven to contain specific types of users that are relevant to security analysts: those who use these platforms as a means of expression of views on personal, political, social and economic issues and those who use these

platforms as an outright means to coordinate security attacks. The operating assumption is that communications are monitored, captured and analysed to detect these trends and infer links to data collected on other dimensions of cyberspace. In this exercise, a set of users (social personas) will fulfil the role of providing social dynamics in the form of discussions that will be integrated into the experimental system.

4.7.3.2 Evidence Identification and Monitoring on the Economic Dimension (Stock Market Crises)

At this stage of the experiment, the researcher seeks to integrate political and/or economic data into the experimental environment. As earlier mentioned in section 2.8.4, types of data from the economic dimension include financial and political trends which can be represented with data such as stock and commodities market data, news articles data etc. To capture the dynamics needed in this scenario, a real-world stock market crash is integrated into the test environment. The timeline of this market crash is adjusted to fit the timeline of the test environment. Donier and Bouchaud (Donier and Bouchaud, 2015) present a plausible explanation for the observed fall in the chosen market prices and tested techniques for analysing these trends. Such claims and methods are taken into consideration in the experimental research environment. In line with the scenario, the stock market price crash represents a crash in delish corporation's company's stocks. This event incorporates the economic dimension into the research experiment and aims to create a link between the crash and events on other dimensions of cyberspace.

In addition to the time-based observed opening and closing prices, the number of shares traded within the security market during the experimental period is also provided. This feature quantifies the significance of changes in stock prices, therefore the significance of the stock crash.

4.7.3.3 Evidence Identification and Monitoring on the Physical Dimension (Data Exfiltration Attack)

The objective of this stage of the experiment is to utilise data from previous stages of the experiment to detect early warning signs of the attack scenario in context. The cyber-attack occurs on the network layer of the physical dimension of cyberspace. To simulate this scenario, a public website, 'delishcorp.com' is setup intended to be accessible to members of the general public (Schreck and Keim, 2013). Sommers et al. (2004) make a case for developing the capacity to generate repeatable, realistic network traffic as being critical for cyber experiments. It is therefore important to incorporate real traffic generation from clients both inside and outside the experiment's network, using available resources. Access logs are made available.

In addition to the public website, the experimental system also simulates a fully functional computer network with acting hosts and services. The computer network is designed with a specific vulnerability for potential exploitation by external adversaries. The network operates under the assumption of a fully functional cyber defence team with monitoring, detection and analytical capabilities. A cyber-attack is conducted in the controlled environment while network traffic logs from all hosts are recorded simultaneously. The network traffic logs are consequently fed as new sources of evidence into the entangled cyberspace model to enhance the predictive capacity of the model.

Furthermore, environmental factors such as weather readings are integrated into the experimental system to model the real-world layer of the physical dimension of cyberspace. The timeline of the weather readings is in sync with the social forum feeds. Finally, the operational theory is a predictive

link between these identified features. The operating network system is designed with a defensive-offensive approach in mind. Certain hosts will play the role of 'nodes in an operational network' while a single host will play the role of an external adversary. Within the operational network, defenders are assumed to be capable of providing defence capabilities to mitigate cyber-attacks. The adversary (the attacker), in this experiment, is in charge of penetrating and challenging the operational network.

The data that represents the scenario described above is presented by IEEE Vast 2011 challenge. The scenario used is designed to mirror the experimental network environment like described in (Schreck and Keim, 2013).

The timelines for each event and the integration of activities on the various layers across this timeline is further explained in chapter 5 of this thesis.

4.7.3.3.1 Network System Components

The development of the network testing environment requires tools and devices similar to those used in complex real-world networks. This research simulates these network behaviours in a virtual environment which is considered cost-effective and time-saving. The virtualization of hosts in an operational network provides a controlled environment for simulating real-world cyber-attack scenarios with minimal interaction with the external network. This approach has the advantage of addressing issues such as technological costs, time and legalities in cyberspace. To achieve the required features with the needed capabilities, the following components were integrated.

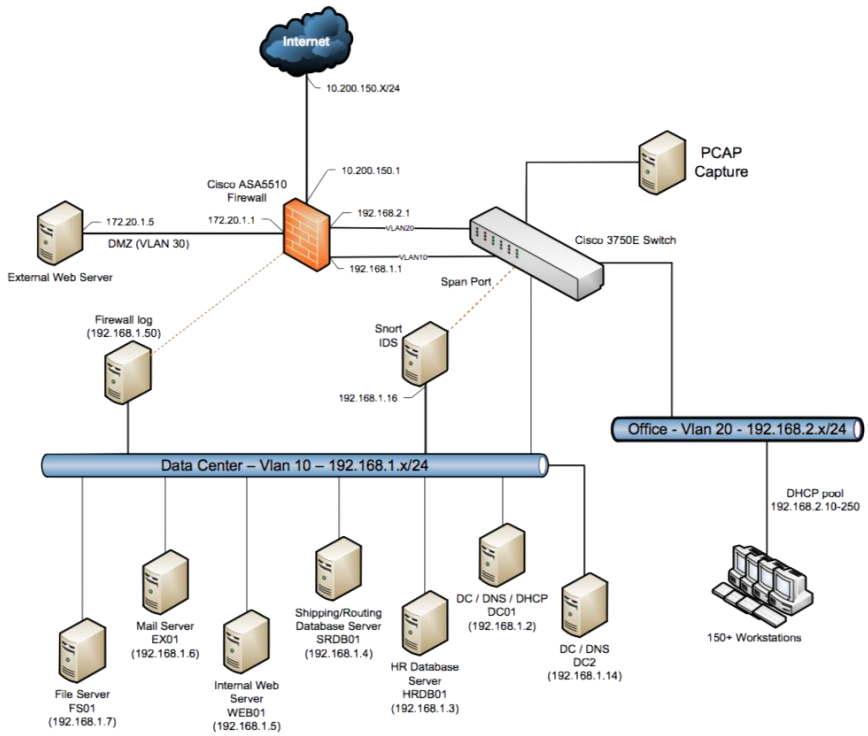

Figure 4-1: Scenario-Based Operating Network Architecture

The figure above is an illustration of fresh delish network architecture with the following notations being associated with the network architecture.

- The subnet domain for the network operating environment is 192.168.1.x/24.

- An IP address (Internet Protocol Address) 'x', described as belonging to the subnet domain indicates that 'x' can be any number between 1-255.
- IP addresses describe as 192.168.2.10-250 indicates the IP address ranges from 192.168.2.10 through 192.168.2.250.

The table below outlines the nodes on the network with detailed descriptions of their functions in the network.

SN	IP Address	Node Type	Description
1	192.168.2.10-250	Office Workstations	Individual workstations of staff.
2	192.168.1.50	Firewall Log	Server that captures system firewall logs.
3	192.168.1.14	DC /DNS Server	Server Running Critical Network Operations.
4	192.168.1.7	File Server	Server holding shared files used by employees.
5	192.168.1.6	Mail Server	The server that stores and routes all email that flows into or out of the network.
6	192.168.1.5	Internal Web Server	Th Server that hosts the company's corporate intranet, including company news site and policy and procedure.
7	192.168.1.4	Routing Database Server	The server containing customer data, including shipping requests and routing information.
8	192.168.1.3	HR Database Server	The Server running the database for employee payroll and benefits.
9	192.168.1.2	DC / DNS /DHCP Server	Server running critical network operations.
10	192.168.2.1	Firewall	Firewall interface to office VLAN.
11	192.168.1.1	Firewall	Firewall interface to data centre VLAN.
12	192.168.1.16	IDS System	Snort IDS interface to the network.
13	10.200.150.1	External Web Server	A web server which hosts the company's external website.
14	172.20.1.1	Firewall	Firewall interface to External Web Server.
15	172.20.1.5	Firewall	Firewall Interface to the internet.

Table 4-6: Network System Components

The following are common ports and services in the experimental network.

- Port 80 – Non-Secured Web Traffic
- Port 443 – Secured Web Traffic
- Port 53: Domain Name Service
- Port 25: Email Traffic

4.7.3.4 Lab Experiments for generation of Network Data

Section 4.7.3.4 outlines the procedure for creating the experimental network. The network was developed for this research with all the requirements as described in the experimental scenario. The system was exposed to cyber experts to ensure it simulated real network behaviour. In this research, it has been established that the main objective is to develop a system with monitoring and data capture capabilities in cyber-attack scenarios.
Additionally, a theoretical foundation is also an important component of this development as it explains what artefacts are needed and what role they play towards the delivery of the solution. The development of the test environment is therefore based on the entangled cyberspace theory discussed in section 3.4. Consequently, the process model shown in **Error! Reference source not found.** is designed to implement the theory of the entangled cyberspace.

4.7.4 Experiment Development

A phased hypothesis testing approach is adopted in delivering the objectives of the experimental design. The experiment is conducted in eight phases, each representing the stages of the social-physical-economic cyber-attack kill-chain presented in section 2.8.6.1 through section 2.8.6.5. Following the outline of the theoretical research model in section 3.4 which is based on the social-physical-economic cyber-kill chain in section 2.8.6, the researcher forms the following hypothesis to test the conceptual model at each stage of the kill-chain.

4.7.5 Experiment Sequence

The experiment has been designed based on the system requirements discussed earlier in section 4.7.2. The experimental sequence design relies on existing literature in areas of social network analysis, cyber kill-chain analysis, ethical hacking, networking and predictive analytics. Each stage of the experiment is set up to represent the corresponding kill-chain phase. The experiment begins at the first phase of the Social-Physical-Economic cyber kill chain (section 2.8.6); the antecedents. Data available from the economic dimension of cyberspace as postulated in the literature review in section 2.8.5.3, which are characteristic of antecedents to cyber-attacks, is used as explanatory variables for the next phase of the kill-chain; the reconnaissance phase, in a structural predictive model. The experiment follows this pattern until the last phase of the kill chain; the attack phase. The experiment is designed to simulate the occurrence of the three main events described in the scenario, sequentially on the kill-chain and across the dimension of cyberspace. The research scenario development relies on these events to produce a real-world situation of inter-connectedness in cyberspace and its usefulness in pre-empting cyber-attacks. In this regard, the experiment is broken down into seven stages.
Each stage of the experiment is designed to address the corresponding research hypothesis by introducing new sources of evidence, from a new dimension into the testing environment. Figure 4-2 shows the entangled cyberspace experiment that summarises the stages of the experiment, starting with identifying the antecedents in the experimental environment. Additionally, there is a provision for dynamic data integration, feature identification and feature selection at each stage of the experiment. Finally, a phased hypothesis testing approach is adopted in delivering the objectives of the experimental design.

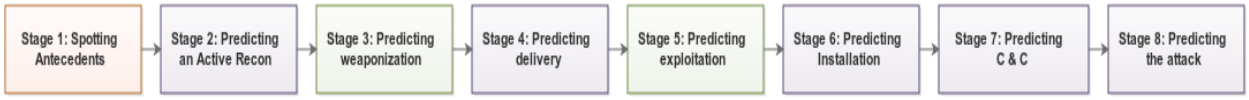

Figure 4-2: Experiment Sequence Diagram

4.7.5.1 Stage 1: Pre-empting Antecedent Activities on the Economic Dimension

Stage 1 of the experiment attempts to create a link between activities on the social dimension of cyberspace and activities on the economic dimension of cyberspace. This stage represents the first event on the Physical-Social-Economic cyber-attack kill-chain to occur on the economic dimension of cyberspace. The researcher builds a predictive model to test for co-integration between variables within and across both dimensions. 'Antecedents' is characterised by events on the social or economic dimensions (Gandhi *et al.*, 2011a) that invoke the cyber-attack kill-chain. The researcher attempts to test the predictive capacity of social activities on economic events with features such as stock market prices and social form discussions. The experiment begins at this stage by testing for co-integrating relationships within the economic and social dimensions and between them. It goes further to construct a predictive model with stock prices as the endogenous variable predicted by a multi-dimensional integration of social and economic dimension features.

4.7.5.2 Stage 2: Pre-empting and Active Reconnaissance on the Physical Dimension

Stage 2 of the experiment attempts to create a link between activities on the economic dimension of cyberspace and activities on the physical dimension of cyberspace. The aim is to build a model that actively predicts an active reconnaissance on the network layer of the physical dimension. To do this, the researcher builds a predictive model to test for co-integration between variables within and across both dimensions. The researcher attempts to test the predictive capacity of economic activities on the network layer with features such as stock prices and network flow data. The experiment begins at this stage by testing for co-integrating relationships within the economic and physical dimensions and between them. It goes further to construct a predictive model with network flow features indicative of an active reconnaissance (Bailey Lee, Roedel and Silenok, 2003; Kinable, 2008) as the endogenous variable predicted by a multi-dimensional integration of physical and economic dimension features.

4.7.5.3 Stage 3: Pre-empting Cyber-Attack Weaponization Phase on the Social Dimension

Stage 3 of the experiment attempts to create a link between activities on the social dimension of cyberspace and activities on the physical dimension of cyberspace. This stage of the experiment assumes that the proliferation of cyber weapons in cyberspace can be detected on the social dimension (Hernández *et al.*, 2016). Using features on the social dimension to represent the 'in progress creation of a cyber weapon', the aim is to build a model that actively predicts this proliferation on the social dimension. To do this, the researcher builds a predictive model to test for co-integration between variables within and across both dimensions. The researcher attempts to test the predictive capacity of physical activities on the social dimension with features such as 'cyber-relatedness' and polarity of social forum discussions. The experiment begins at this stage by testing for co-integrating relationships within the social and physical dimensions and between them. It goes further to construct a predictive model with social forum features indicative of a weaponisation phase as the endogenous variable predicted by a multi-dimensional integration of physical and social dimension features.

4.7.5.4 Stage 4: Pre-empting the Cyber Delivery Phase on the Physical Dimension

Stage 4 of the experiment attempts to create a link between activities on the social dimension of cyberspace and activities on the physical dimension of cyberspace. This stage of the experiment

assumes that the ability of theoretical model to identify a cyber weapon delivery within a computer network. With a combination of features from the social and physical layers, the researcher attempts to investigate co-integrating relationships between activities on both layers. Using features on the physical dimension to represent the injection of a cyber weapon into a victim's network, the aim is to build a model that actively predict this event with features from other dimensions of cyberspace. To do this, the researcher builds a predictive model to test for co-integration between variables within and across both dimensions. The researcher attempts to test the predictive capacity of physical and social activities on the physical dimension with features such as derived from network flow data (Sans Institute, 2012; Almukaynizi *et al.*, 2017). The experiment begins at this stage by testing for co-integrating relationships within the social and physical dimensions and between them.

4.7.5.5 Stage 5: Pre-empting the Exploitation Phase on the Physical Dimension

Stage 5 of the experiment attempts to predict a cyber exploitation in a victim's network. At this stage of the experiment, the researcher combines data from three dimensions of cyberspace in a single model. The previous stage of the experiment dealt with data on the physical dimension as endogenous variables. This stage assumes a co-integrating relationship between data on the physical dimension and/or other dimensions of cyberspace. At this stage, we incorporate data from the three identified dimensions of cyberspace and apply a feature selection algorithm (Bonev, Escolano and Cazorla, 2008) to filter the relevant exogenous variables for the system. The endogenous variable is derived by methods described in (Khater and Overill, 2015) to represent the actual execution of exploit injected into the victim's network in the previous stage.

4.7.5.6 Stage 6: Pre-empting the C&C Phase on the Physical Dimension

Stage 6 of the experiment attempts to predict the C&C phase of the cyber kill chain by actively identifying the activities of botnets in the victim's network. The researcher begins by characterising activities of botnets on a victim's network. The characterisation of botnets is achieved using methods defined by (Raghava, Sahgal and Chandna, 2012) for identifying botnets in computer networks. This stage of the experiment assumes that the ability of theoretical model to identify the communications of botnets in a computer network and predict these activities.

Additionally, with a combination of features from the three dimensions of cyberspace identified in this research, the researcher attempts to investigate co-integrating relationships between activities on these dimensions. Using features on the physical dimension to represent the activities of botnets in a victim's network, the aim is to build a model that actively predicts this event with features from other dimensions of cyberspace. To do this, the researcher builds a predictive model to test for co-integration between variables within and across these dimensions. The researcher attempts to test the predictive capacity of physical, social and economic activities on the physical dimension.

4.7.5.7 Stage 7: Pre-empting Of Cyber-Attack on the network layer of cyberspace

Stage 7 of the experiment assumes a cyber network attack on a target's network. The researcher begins by characterising activities of a cyber-attack on the target's network and assumes the ability of theoretical model to identify the activities of a cyber-attack in a computer network and predict these activities. Additionally, with a combination of features from the three dimensions of cyberspace identified in this research, the researcher attempts to investigate co-integrating relationships between activities on these dimensions. Using features on the physical dimension to represent the perpetration of a cyber-attack in the victim's network, the aim is to build a model that actively predicts this event with features from other dimensions of cyberspace. To do this, the researcher builds a predictive model to test for co-integration between variables within and across these dimensions. The researcher attempts to test the predictive capacity of physical, social and economic activities on the physical dimension.

4.8 CONCLUSION

This chapter discussed the methods for research and experimental design. This research aims to investigate techniques to detect early warning signs of a cyber-incident in cyberspace and thus pre-empt cyber-attacks. The events at each stage of the experiment are considered ground truth. This aim stems from an ultimate contribution to increasing the state of cyber situational awareness. This study uses a positivist approaches together with some proven research design methods to access the proven advantages of such areas of research design and natural sciences. This chapter also provided a critical understanding of the experimental design and structure of analysis to be conducted at each phase of the experiment. The study is done in here phases. The first is the scenario development which creates the events around which study data is simulated. These scenarios are modelled to contain the behaviour of interest. The second stage uses proven statistical techniques and lab experiments to generate the data for this research. The third phase builds models to test the proposed theory using simulated data. Co-integration, Vector Auto-regressive and causal models were used to verify the hypothesised theoretical model and test relationships between variables of interest. The outcome is then used to develop a predictive model for active cyber defence. The next chapter provides a practical implementation of the developed theoretical framework using an experiment built around the scenario.

5 CHAPTER 5: DATA ANALYSIS

5.1 INTRODUCTION

This chapter presents the methods for data gathering, data preparation and data analysis in this research. The chapter presents the benchmark data used in validating the theoretical constructs presented in section **Error! Reference source not found.** of this research. Additionally, this phase of the research specifically highlights the analytical techniques applied to the benchmark data. The data preparation follows Hair et al. (2010)'s six-step structured approach to multivariate model construction. The process of building of the time-series prediction model at each stage of the experiment is also outlined, and the proposed analytical framework follows (Kuhn & Johnson Applied predictive modelling) step-by-step guide for building predictive models. Finally, this chapter tests the implementation of the structural analytical theory, vector autoregressive models, designed to test the suggested experimental hypothesis between the underlying constructs of the entangled cyberspace model.

5.2 DATA PREPARATION

This section explains how the data used in this research was acquired, formatted and transformed. Data acquisition, data formatting, data aggregation, data transformation and data validation was performed (Pyle, Editor and Cerra, 1999) in order to ensure data consistency, comprehensiveness and relevance to the problem statement. This study comprises of multiple sources of evidence which are assumed to represent theoretical concepts to be tested. Therefore, datasets were uniquely selected to represent activities on the physical, social and economic dimension respectively. In order to achieve a representative data model, the researcher creates a data space for each dimension of cyberspace represented in the literature review. In this regard, this study works with three data spaces, where each dataspace contains multiple sources of evidence. A 5-stage data preparation approach (Pyle, Editor and Cerra, 1999) was applied to each dataset on each dimension using a functional scripting language[1]. The data preparation strategy used on each data dimension is visualized in the figure below.

[1] All scripts and files used in producing the results seen in this section are available. see Appendix 1.1

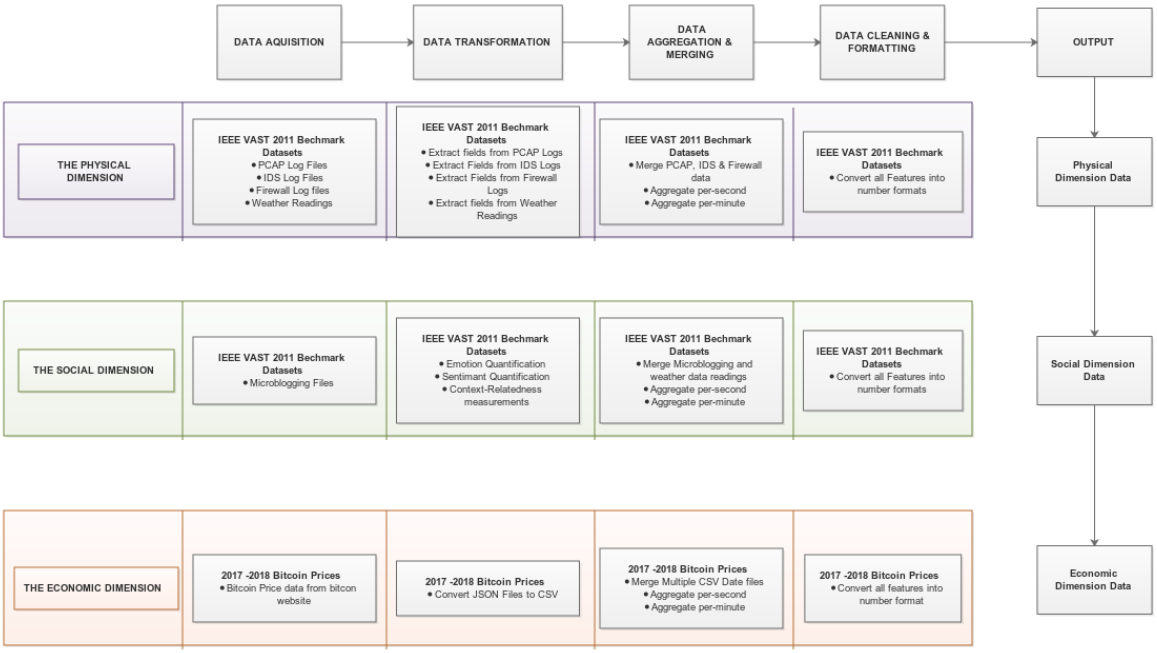

Figure 5-1: Data Preparation Framework

5.2.1 Data Acquisition

This section describes how the data used in this research was collected. It presents the sources of evidence that feeds model testing of the theoretical constructs. This research uses subsets of the 2011 VAST IEEE visual analytics challenge datasets with a combination of real-world datasets simulated from the financial market prices. The IEEE Visual Analytics Benchmark Repository (Cook *et al.*, 2011) contains benchmark datasets and experiments for theoretical testing and development. The ground truth in benchmarks follows the events outlined in the scenario development, allowing for accuracy in metrics comparison and computation. Primarily, benchmark datasets are assumed to contain theoretical constructs under observation, and it is the researcher's aim to use these datasets to test how well the research artefact picks up the nature of entanglements embedded in the datasets. The datasets used in this research are selected specifically to create a uniform representation of each dimension of cyberspace as identified in the literature review[2].

5.2.1.1 Data Acquisition on The Physical Dimension

Two categories of data are made available on the physical dimension; physical network data for a fictitious company and weather data readings for the metropolitan city in the scenario. These datasets represent the network layer and real-world layer of the physical dimension of cyberspace as discussed in the literature review. The datasets on the network layer are provided as evidence for improving the overall cyber situational awareness of the company (Cook *et al.*, 2011). Four of the six datasets provided by the IEEE vast (Cook *et al.*, 2011) are used in this research.

[2] All scripts and files used in producing the results seen in this section are available. see Appendix 1.2

In this research, we aim to incorporate activities on the physical dimension of cyberspace in our theoretical model. The datasets provided on the physical dimension are as follows:

- Network flow data from the network layer.
- Intrusion Detection Logs from the network layer.
- Firewall Logs from the network layer.
- Weather data from the real-world layer.

The scenario developed in this research incorporates log data from the network layer and weather data readings from the real-world layer of cyberspace. Three types of network log datasets were provided: network traffic logs, firewall logs and intrusion detection logs.

Network flow data (PCAP Files) is provided for three days of network observations. The PCAP files are packet header data for all communications between source and destination IP addresses within the company's network. The network log data was collected for the 13th, 14th and 15th of May. Communications are distinguished by the protocol-level information in the header, the source and destination port addresses, service levels and time.

In addition to network flow capture data, network firewall logs are also provided. The firewall for the experiment in this research is a Cisco Adaptive Security Appliance 5510 (Cisco, 2017). There are one or more firewall log files for each of the three days under observation. Each file is a maximum of 512 megabytes, containing both internal network events and external network events. All internal network events involving traffic between VLANs are logged, and all traffic relating to external networks is logged. This data is captured in a comma-separated values (.csv) file, with the most recent entry at the top.

The intrusion detection system used in the experimental design is Snort (Dwivedi and Tripathi, 2015). Snort is a customizable intrusion detection system used to detect different types of network traffic based on a pre-defined set of rules.

This dataset includes one IDS log for each day in text file (*.txt)* format that includes all intrusion detection events for the entire network for that day. Each log entry contains many detailed fields for each detected event. Each event relates to a entry in IDS logs collected by network SNORT. These fields are shown and explained in the table below:

S/N	CONTENT	DESCRIPTION
1	[**] [*Snort Rule File*] *Text of specific rule violated* [**]	Text of specific rule violated
2	[Classification: an optional generalized description of the alert]	Classification (if present)
3	[Priority of the alert]	Priority
4	Date/Time MM/DD-HH24:MI:SS.milliseconds	Date/Time
5	Source IP/Port	Source IP/Port
6	Destination IP/Port	Destination IP/Port
Remaining lines	*No content used in the challenge*	None

Table 5-1: Intrusion Detection Fields

In addition to the network layer, the real-world layer is also represented on the physical dimension.

Weather data readings are provided for the metropolitan city in the scenario environment. The data was recorded at specific weather stations daily for a period of 20 days. The daily average wind speed, overall daily weather classification and wind direction are recorded in a .csv file
A firewall is set up to monitor all incoming and outgoing traffic on a network and additionally filters specific types of connections from entering a network. In addition to IDS logs and network traffic logs, firewall logs are also provided from the experimental environment. The firewall logs record every connection made to the target network. Additionally, the firewall log records if each connection was rejected or accepted within the number of inbound connections and outbound connections on a network.

SN	FEATURE NAME	FEATURE DESCRIPTION	REFERENCE
1	**Date Time**	Date and Time when the activity was performed	(Fu *et al.*, 2009)
2	**Syslog Priority**	Priority of the log message	(Fu *et al.*, 2009)
3	**Operation**	Type of activity being performed	(Fu *et al.*, 2009)
4	**Message Code**	Message Code	(Fu *et al.*, 2009)
5	**Protocol**	Connection Protocol Type	(Fu *et al.*, 2009)
6	**Source IP**	Source IP associated with the activity	(Fu *et al.*, 2009)
7	**Destination IP**	Destination IP associated with the activity	(Fu *et al.*, 2009)
8	**Source Hostname**	Source Hostname associated with the source IP	(Fu *et al.*, 2009)
9	**Destination Hostname**	Destination Hostname associated with the destination IP	(Fu *et al.*, 2009)
10	**Destination Port**	Destination port associated with the destination IP	(Fu *et al.*, 2009)
11	**Source Port**	Source port associated with the source IP	(Fu *et al.*, 2009)
12	**Destination Service**	Name of service associated with the destination port	(Fu *et al.*, 2009)
13	**Connections Built**	Number of connections built in this operation	(Fu *et al.*, 2009)
14	**Connections Torn**	Number of connections torn down in this operation	(Fu *et al.*, 2009)

Table 5-2: Firewall Fields

5.2.1.2 Data Acquisition on the Social Dimension

Similarly, the researcher aims to incorporate data from the social dimension of cyberspace. Data on the social dimension of cyberspace are generated by cyber-personas, behind which exists real-world personas. The thoughts and social perception of these cyber-personas are captured as text data, more specifically social forum discussions. The social dimension of cyberspace characterizes activities from cyber personas and real-world personas. This dimension seeks to capture interactions, thoughts, perception and social interactions between personas in cyberspace. In order to incorporate the concepts of the persona and cyber persona in this research, social interactions from microblogging

platforms are provided. This experiment represents these interactions with the text feeds provided by the IEEE vast data. Within the scenario, the microblogging messages were collected from various devices with GPS tracking capabilities. These devices include laptop computers, handheld computers, and cellular phones. The feeds are also tagged with unique identifiers for users posting these messages at any given time. The initial variables provided are outlined in the table below:

SN	VARIABLE NAME	VARIABLE DESCRIPTION
1	ID	Personal Identifier of the individual posting the message.
2	Created_at	Date and Time the message was posted.
3	Location	Geographical coordinates of the device at the time of posting.
4	Text	The message that was posted.

Table 5-3: Initial Features from Microblogging Feeds

This data is the mini-challenge 1 data of the 2011 IEEE Vast visual analytics challenge available at (HCIL, 2013). 1,023,077 chat text was collected from 73,928 users in the chatroom over a 20-day time frame between the 30th of April and the 20th of May.

5.2.1.3 Data Acquisition on the Economic Dimension

The economic dimension contains evidence that supports the effects of economic, political and cultural events on cyber-events. Data collection on the economic layer used a real-world simulation of the desired effect. The scenario development in this research requires the representation of a company's stock price crash. For this challenge, the researcher gathered data from the 2017 bitcoin price crash. The Bitcoin (BCH) price crash of 2017 took place from July 2017 after maintaining a steady increasing trend all year round. The per-minute figures for the opening, closing, market high, market low, market capitalization and price volume was collected for each day for the year 2017. The data was collected for a one-year window following the start of the increasing trend at the start of the year through the crash up until the end of the year. This duration was selected for two reasons. First, the experiment requires that a base of "normality" be established. Incorporating a time period where prices followed a "normal" trend helps the research artefact to identify sudden changes in this trend. Second, existing research indicates that long-term analysis of financial markets provides a more accurate overview of the long-term performance or sustainability of financial prices (Jin and An, 2015). The table below shows the description of the variables used on this dimension.

SN	**Variable Name**	**Variable Description**
1	Open	Per Minute Opening Prices for stock prices.
2	Close	Per Minute Closing Prices for stock prices.
3	High	Highest Stock Price recorded per minute.
4	Low	Lowest stock price recorded per minute.
5	Volume_USD	An average number of Company's shares traded within a given time frame.
6	Market Capitalization	Average total recorded value of Company's stocks within a given time frame.

Table 5-4: Initial Features from Stock Data

The rationale behind including this real-world data into the experimental design is to ensure that all dimensions are represented in the experiment. These prices are to reflect the effects of activities on the

economic layer on events in other dimensions of the information space. The prices are included in the experiment to determine the possibilities of our proposed approach, picking up these evidences if they exist. Additionally, this research is concerned with proving that design artefacts are capable of picking up entanglements within various dimensions in cyberspace if they exist. It is geared toward establishing prove for an inter-dependent relationship between entities and elements in cyberspace.

5.2.2 Data Transformation

This section describes how the data collected in this research was transformed. Data transformation is necessary to extract useable features from data in a way that dynamics and relationships contained in the data are retained. Datasets provided on each dimension are transformed into suitable formats for predictive modelling and result analysis. Useful features are derived from existing data on each dimension using well-tested techniques that retain the dynamics of relationships between features in the datasets. Features are extracted to represent each event in the experimental design using transformation and quantification techniques.

5.2.2.1 Data Transformation on Physical Dimension

On the physical dimension, data transformation is performed on the three datasets provided, network traffic logs, intrusion detection logs and firewall logs.

5.2.2.1.1 Network Traffic Fields

Fields from the network traffic logs, in the pcap formats, are extracted using the tshark library on a Linux command line interface. The following fields were extracted from each packet in the pcap files.

SN	VARIABLE NAME	VARIABLE DESCRIPTION	REFERENCE
1	**Source IP**	Internet Protocol Address of the device sending the IP Packet.	(Borja, 2013; Sanders, 2017)
2	**Destination IP**	Internet Protocol Address of the device receiving the IP Packet.	Borja, 2013; Sanders, 2017)
3	**Source Port**	Port number from which the data or request should be sent from the remote host.	Borja, 2013; Sanders, 2017)
4	**Destination Port**	Port number to which the data or request should be sent on the receiving host.	Borja, 2013; Sanders, 2017)
5	**Packet Arrival Time**	The timestamp when the packet was received.	Borja, 2013; Sanders, 2017)
6	**Epoch Time**	Also known as “UNIX” time is the number of seconds since January 1st 1970.	Borja, 2013; Sanders, 2017)
7	**IP Protocol**	The protocol used for session communication. Can be 1 for UDP packets, 4 for IPV4 packets, 17 for UDP packets or 6 for TCP packets.	Borja, 2013; Sanders, 2017)
8	**Total Length**	The packet length in bytes	Borja, 2013; Sanders, 2017)
9	**Time To Live**	The remaining lifetime of a packet when it is floating in a network.	Borja, 2013; Sanders, 2017)
10	**IP Flags**	Flags field used to control how a specific IP packet is treated by a device	Borja, 2013; Sanders, 2017)
11	**Entropy of Source**	Ratio of packets with a given source IP	(Shah and Tanvi, 2006)

	IP	to ratio of all Source IPs within a given time frame.	
12	**Entropy of Destination IP**	Ratio of packets with destination IP to ratio of all packets receiving IPs within a given time frame.	(Shah and Tanvi, 2006)
13	**Entropy of Source Port**	Ratio of packets with a given source port to ratio of all Source ports within a given time frame.	(Shah and Tanvi, 2006)
14	**Entropy of Destination Port**	Ratio of packets with a given destination port to ratio of all destination ports within a given time frame.	(Shah and Tanvi, 2006)
15	**Packet Type**	Protocol used to transmit data over the network.	(Shah and Tanvi, 2006)

Table 5-5: Features Extracted from Network Traffic Logs

The twenty-one extracted fields were converted into a CSV file where each record in the CSV file represents filed data for a single packet identified by the packet number.

5.2.2.1.2 Intrusion Detection Log Fields

A plain text file with network intrusion details is provided for each day between the 13th and 15th of May. Each text file contains log entries for event caught by snort. Each plain text file was transformed into a list of events where each item in the list is an event caught by snort recorded in the file. For each log entry, the researcher extracts all available variables such as the snort rule violated, alert classification, the priority of the alert, destination ports, source ports, destination IP, source IP, date and time of the alert. The table below outlines the features extracted from the Intrusion detection system logs.

SN	FEATURE NAME	FEATURE DEASCRIPTION	REFERENCE
1	**Date of Event**	Date of Alert	(Fu *et al.*, 2009)
2	**Time of Event**	Time of Alert	(Fu *et al.*, 2009)
3	**Alert Classification**	Description of Alert, e.g. portscan. Pingsweep, Attempted Denial Of Service	(Fu *et al.*, 2009)
4	**Priority of Alert**	Priority Level of alert	(Fu *et al.*, 2009)
5	**Destination IP**	Destination IP of alert	(Fu *et al.*, 2009)
6	**Destination Port**	Destination port associated with destination IP	(Fu *et al.*, 2009)
7	**Source IP**	Source IP of Alert	(Fu *et al.*, 2009)
8	**Source Port**	Source port associated with source IP	(Fu *et al.*, 2009)

Table 5-6: Features Extracted from Intrusion Detection Logs

5.2.2.1.3 Firewall Log Fields

The firewall data was provided in a CSV format therefore little, or no data transformation was needed at this stage. However, selected features would be extracted from this dataset at a later stage of data preparation.

5.2.2.2 Data Transformation on the Social Dimension

In order to ensure a uniform data format for all datasets across three dimensions, the researcher applied data transformation techniques that quantify the given text data while retaining as much useful information as possible. The original data was aggregated on time to produce a per-minute summary of the original data. This first stage of data transformation produced a new dataset with four variables: the timestamp of the current minute, a combination of all messages posted within that minute, the total number of users posting within the minute and the total number of unique places recorded by the geographical location.

The next phase of data transformation on the social dimension involves the quantification of the messages posted in the chat forum. This quantification phase follows a four-part approach: quantification based on opinions, quantification based on context-relatedness, quantification based on the uncertainty of information and quantification based on plutchnik's wheel of emotions (Plutchik, 1982).

5.2.2.2.1 Quantification of Microblogging Feeds

This subsection outlines the methods used in the quantification of microblogging feeds on the social dimension. The quantification techniques presented here are taken from an extensive review and evaluations of methods used in contemporary literature. The algorithms and techniques critically discussed in the literature review in (Plutchik, 1982; Nigam, Lafferty and Mccallum, 1999; Rose *et al.*, 2010; Calefato, Lanubile and Novielli, 2017)are implemented in this section and applied to microblog data. The quantification of microblogging data is important to extract information from text data in a format that fits into predictive models.

5.2.2.2.1.1 Word Count

In modern English language, individual words are separated by an empty space. Therefore, given a text corpus ***C***, the word count ***WC*** is defined as the number of individual words in ***C*** delimited by an empty space:

$$WC\ (C) = \sum_{i=1}^{n} [C_i = 1]$$

Equation 5-1: Word Count In Text Data

However, the relevance of the measure of "word count" is based on the pre-assumption of English language texts. In addition, the word count is also used as a normalizing feature for other frequency-based quantification of text data.

Each text in the microblogging feed is first split on a specific delimiter (in the English language, a space to indicate the separation of words) to create an array ***F*** of words. The word count is then given as the length of ***F***.

5.2.2.2.1.2 Opinion mining

Given a text corpus ***C***, we a five-item tuple "Opinion" is derived containing the positivity, negativity, neutrality, sentiment orientation, sentiment intensity of C.

Opinion (C) = (PS, NG, NE, SO, SUB)

Where;

PS: The degree of positivity of C
NG: The degree of negativity of C
NE: The degree of neutrality of C
SO: the Sentiment orientation or polarity of C
SUB: The opinion subjectivity of C i.e. the measure of bias expressed in C.

The general idea behind sentiment analysis as discussed in chapter 2, is to calculate an opinion score for a given text corpus. Here, the researcher represents sentiment as a tuple with three items, sentiment orientation (***SO***), sentiment ratings (***SR***) and sentiment intensity (***SI***):

Sentiment = (SO, SR, SI)

Calculating the sentiment orientation involves assigning a measure of positivity, negativity or neutrality to the sentiment score ***SS*** of a given text corpus. A threshold or boundary of neutrality is then required to classify text as positive, negative or neutral. The point of neutrality, usually the threshold, is the point at which it is assumed that no opinion is expressed in the given text. This point is usually 0. In its most standard form, the sentiment score of a given text corpus SS is given as:

$$SS = \sum_{i=1}^{N}[P_n = 1] - \sum_{i=1}^{N}[N_n = 1]$$

Equation 5-2: Sentiment Score of Text Data

Where *P* is the number of positive words in the text corpus and *N* is the total number of negative words in the document. In this case, the threshold of neutrality is 0 which indicates the absence of sentiment in an opinion. If ***SS*** is less than 0, the text corpus is said to have an overall negative score. On the other hand, if the text corpus is greater than 0, the text corpus is said to have an overall positive score. In addition, this method uses opinion lexicon of positive and negative words designed by (Lui, 2015) to tag each word in a given text corpus as "positive" or "negative" after appropriate data cleaning techniques have been applied to the text corpus.

Sentiment intensity refers to the strength of the opinion having been classified as positive, negative or neutral. An opinion can be weakly positive, weakly negative, strongly positive or strongly negative.

5.2.2.2.1.3 Entropy

Let us consider a binary variable X, which is associated with precisely a single word. X = 1 if the word exists in a given document and X = 0 otherwise. The entropy of the entropy of the word X is given as:

$$H(X) = -P(X = 0)log_2\,P(X = 0) - P(X = 1)log_2 P(X = 1)$$

Equation 5-3: Shannon's Entropy of Text Data

The negative logarithm of each word in each text in the microblog feeds is summed up to produce a quantified measure of information uncertainty for each text. In information theory, the degree of randomness of a variable is directly correlated with the entropy and in contrast the greater the certainty or order of the variable, the smaller the entropy.

The above definition of entropy indicates that the more randomness is associated with a specific variable of interest the bigger the uncertainty associated with it as well. Using entropy measurements for language texts involves estimating the measure of credibility of the information transmitted by a communication channel. When estimating the information entropy of a language text which consists of letters, words and sentences, each letter occurrence is treated as a sequential realization of a certain pattern. The amount of information contributed by the occurrence of each letter is therefore dependent on the intended sequence that forms the entire language text. Therefore, a probability distribution of all letters and all possible sequential combinations of letters are considered. Using this technique reveals the usefulness of each letter or word to the expression of the entire language text and also the level of complexity of the language text. For example, a long string of language text, containing a single letter, 'aaaaaaaa' would have an entropy of 0 indicating an absence of useful information in the text. This application does not account for covert channels of communication especially in areas of tracking cyber-related activities embedded in unstructured data [ref] the methods described in the methodology are limited to normal channels of communications as standard techniques in research would apply. Whilst it is possible to have these covert channels of communication, the approach works on picking up signatures within unstructured data as is, i.e it does not assume any prior structure to the data or signatures within the language used.

5.2.2.2.1.4 Plunik's Emotion Detection

The system used for classification of text into the eight dimensions of Plutchnik's emotions is based on gold standard datasets provided by (Norcross, Guadagnoli and Prochaska, 1984). Our approach trains a multi-class classification model using gold standard labelled dataset. Each text in the gold standard dataset is labelled as either one of Plutchnik's eight dimensions of emotions. The application of trained recurrent neural network (RNN) models produces eight new probabilities for each text in microblog feeds: Anger, Anticipation, Sadness, Trust, Joy, Fear, Disgust and Surprise.

5.2.2.2.1.5 Context Relatedness

This section explains how the lexicon used for quantifying the various events of interest in the microblogging feeds was created. Given a set of text corpus C, that represents discussions on a particular topic of interest, the researcher seeks to create a wordlist of K terms that are seen to frequently occur across all corpus samples. Working under the assumption that these terms captures the event of interest, the researcher quantifies microblogging feeds combining the relative frequency of occurrence of these terms in each feed and pairwise mutual information of occurrence with other words in the feed.

Problem Definition

Given a corpus C and a wordlist D, where d_1 to d_n represents a set of words that are frequently used during the occurrence of the event of interest, a rank K is derived where K is the number of matches of d_i in C. Here we say the value K represents the "context-relatedness" of the corpus C.

- ***Creating the wordlists***

The experiment follows three different events on the social dimension: Discussions about a cyber-attack, discussions about a fire incident and discussions about flu symptoms to signify the dispersion of a flu epidemic. Additionally, the researcher uses the cyber-attack discussions to characterize the weaponization phase of the kill chain in the experiment. In order to accomplish these tasks, the researcher creates a wordlist for each event listed above, a flu event, a cyber-incident, a fire incident and cyber weaponization.

The 'flu wordlist' was created by targeting words that signified symptoms of a flu. These words were taken from flu descriptions from patients and medical experts on medical sites such as cdc.gov (cdc.gov, 2016), nhs.uk (Hern and Gibbs, 2017) and webmd.com. Similarly, the 'fire wordlist', is created by extracting discussions, complaints and alerts from incident response handles on twitter such as @UkFireServices, @ChicagoFireServices, @OttFire and @911FireUK. The researcher was careful to only collect tweets from other users mentioning these handles as opposed to collecting tweets posted by the handles themselves.

In both scenarios, in order to create a wordlist of relative terms, the researcher applies keyword extraction technique based on terms that best describe the subject of each text provided. The keyword extraction technique is based on the rapid automatic keyword extraction described by (Rose *et al.*, 2010). The RAKE algorithm begins keyword extraction by creating a set of candidate keywords from the text corpus. To create a set of candidate keywords, the algorithm starts off by splitting the text into an array of words using specified delimiters. This array of words is then grouped into contiguous words using specific words and phrases as delimiters. Words occurring together within groups are considered to be candidate keywords. Each identified word is then given a score based on the frequency of occurrence and degree of co-occurrence with other words. Finally, the top K scoring terms are selected as the extracted keywords from the corpus.

To create the 'cyber-incident' wordlist, cyber discussions from known cyber-incident related twitter handles, discussions from specific subreddits and blog posts from cyber-attack reporting blogs such as cyberwarnews.info, hackreads.com and hackmaggedon.com were used. Similarly, in creating the 'weaponization' wordlist, exploit and vulnerability descriptions from exploit database, google hacking database and common vulnerabilities database (CVE) was used.

In the cases of the 'cyber-incident' and 'weaponization' wordlists, the RAKE keyword extraction algorithm was extended to include a pointwise mutual information score for each co-occurrence of words. The researcher creates a PMI matrix of candidate terms, and the top K terms were selected. Where $K = \frac{1}{3} * N$ and N is the total number of words in the text corpus.

To calculate the degree of 'relatedness' of a document, we rely on a word frequency-based algorithm that is a function of the document's words and our generic wordlists. For the purpose of this evaluation, our estimate of the 'relatedness' of a given text is a measure of the summed term scores of individuals words in the text matched to terms in our generic wordlist. The term scores in the wordlist are re-scaled on a scale of 0 through 100. The relatedness of a text is therefore estimated as the average of the term scores of all context-related terms in the given text. Given that all term scores were previously put on a scale of 0 through 100, the expected estimate for the relatedness of any given text should also be on a similar scale. Therefore, these estimates can be represented as a percentage. Note that each word in a text is scored regardless of its frequency of occurrence; therefore, a word $\boldsymbol{W}$ with a word frequency of 3 will add $(W_{apmis}) * 3$ to the total text score.

1.1.1.1 Data Transformation on the Economic Dimension

Data on the economic dimension consists of json files for every minute of every day for the time period December 31st 2016 to January 1st 2018. As earlier mentioned, the rationale behind including this data into the experimental environment is to simulate a stock market crash for the fictitious companies. Bitcoin prices were reported to have experienced a massive crash in the third and fourth quarters of 2017 (CoinMarketCap, 2018). This real-world simulation ensures that the data used captures the dynamics of analytical interest intended. Each json file was converted into csv representing the variables shown in Table 5-4 above. All resulting csv files were further

chronologically merged into a single csv file. The raw data is available at (CoinMarketCap, 2018). Using the method stated, the researcher collected the raw data for everyday within the time frame in json format. This resulted in 396 json files. Each json file contained with each object representing a tuple of seven items (Opening Price, Closing Price, High, Low, Volume BTC, Volume USD, Market Capitalization). The resulting list of tuples contained 1440 tuples each represent a single minute of the given day. The json file for each day was converted into a .csv format using only unique records (all duplicate records were discarded). All 396 csv files were combined into a single file and sorted chronologically. All variables within the final csv file were converted into their appropriate format and the final data represents the (Date, Time, Opening Prices, Closing Prices, Highest observed price within the minute, Low observed price within the minute, Volume USD and Market Capitalization) for every minute of every day within the time frame. The table below shows a summary of the variables extracted on this dimension.

5.2.3 Data Merging and Aggregation (Generating Time Series)

This study uses simple aggregation analysis to generate time signals for predictive model development. At this stage of the data preparation, it was important to integrate all datasets on a similar cohesive timeline for analysis. On the physical dimension, individual fields in the packet data, IDS log data and firewall data are analysed to extract useful information for effective time series analysis. The table below outline the variables derived from a time-based aggregation analysis and combination of features from each data source on the network layer.

	FEATURE NAME	FEATURE DESCRIPTION	DATA SORCE
1	**Total Number of Packets**	The total number of packets observed for a given time window T.	Pcap Logs
2	**Total Number of Source IPs recorded**	The total number of unique source IPs observed for a given time window T.	Pcap Logs
3	**Total Number of Destination IPs recorded**	The total number of unique destination IPs observed for a given time window T	Pcap Logs
4	**Total Number of Source Ports recorded.**	The total number of unique source ports observed for a given time window T	Pcap Logs
5	**Total Number of Destination Ports recorded.**	The total number of unique destination ports observed for a given time window T	Pcap Logs
6	**Mean Epoch Time**	Mean of all epoch times within a given window T.	Pcap Logs
7	**Mean Ip Length**		Pcap Logs
8	**Mean Transmitted Bytes**		Pcap Logs
9	**Flag**	The total number of flags raised within a given time window T. We expect an equal variance in the long term for normal traffic	Pcap Logs

		taking the total number of connections into consideration.	
10	**Mean Frame Length**		Pcap Logs
11	**Mean TTL**		Pcap Logs
12	**Network Average Hold Time**		
13	**Network Traffic**	The amount of data moving through the network for a given time window T.	
14	**Network Congestion**	The amount of people communicating on a network within a given time window T.	
15	**Number of total connections made**	The total number of inbound and outbound connections made within a given time window T.	Firewall Logs
16	**Total number of packets sent**	The total number of outbound packets sent within a given time window T.	Pcap Logs
17	**Total number of inbound connections**	The total number of inbound packets received within a given time window T.	Firewall Logs
18	**Total number of outbound connections**	Total number of outbound connections made within a given time window T.	Firewall Logs
19	**Total number of connections built**	The total number of inbound connections made within a given time window T.	Firewall Logs
20	**Total number of connections torn down**	The total number of connections rejected by firewall within a given time window T.	Firewall Logs
21	**Total Number of services**	Total Number of services running on destination ports within a given time window T.	Firewall Logs
22	**Total Number of Operation**	Total Number of operations / activities performed within a given time window T.	Firewall Logs
23	**Average Syslog Priority**	The average level of priority for alerts received within a given time window T.	Firewall Logs
24	**Total number of IDS alerts**	The total number of alerts raised by the Intrusion detection system within a given time window T.	IDS Logs
25	**Average level of intrusion priority**	The average level of Intrusion detection priority within a given time window T.	IDS Logs
26	**DOS Alerts**	Total number of Attempted denial	IDS Logs

		of service alerts raised within a given time window T.	
27	**Port Scan Alerts**	Total number of port scan alerts raised within a given time window T.	IDS Logs
28	**Ping sweep Alerts**	Total number of pingsweep alerts raised within a given time window T.	IDS Logs
29	**Spp_frag3**	Total number of spp_frag3 alerts raised within a given time window T.	IDS Logs
30	**Total number of DNS requests**	The total number of DNS requests made on port 53 within a given time window T.	Pcap Logs
31	**Total number of UDP Connections**	The total number of UDP Packets transmitted within a given time window T.	Pcap Logs
31	**Total number of TCP Connections**	The total number of TCP Packets transmitted within a given time window T.	Pcap Logs
31	**Total number of ICMP Connections**	The total number of ICMP Packets transmitted within a given time window T.	Pcap Logs
32	**Total Number of HTTPS requests**	The total number of HTTPS requests made on port 443 within a given time window T.	Pcap Logs
33	**Total Number of Email Requests**	The total number of EMAIL requests made on port 25 within a given time window T.	Pcap Logs

Table 5-7: Network Data Aggregated Features

These features were generated from simple per-second aggregates of previous extracted features from the three datasets provided on this dimension. Similarly, on the social dimension, all probability features were merged with newly quantified text information to produce the features listed in the table below:

SN	Derived Variable	Description	Reference
1	**Word Count**	The number of unique words in a given text document.	(Blumenstock, 2008)
2	**Opinion: Sentiment - Positivity**	Degree of the positive emotional effect or positive polarity expressed in a given text document.	(Pang and Lee, 2006; Cambria *et al.*, 2013; Lui, 2015)
3	**Opinion: Sentiment - Negativity**	Degree of the negative emotional effect or negative polarity expressed in a given text document.	(Pang and Lee, 2006; Cambria *et al.*, 2013; Lui, 2015)
4	**Opinion: Sentiment -**	Degree of unbias expressed in	(Pang and Lee, 2006;

	Neutrality	a given text document.	Cambria *et al.*, 2013; Lui, 2015)
5	**Opinion: Sentiment - Subjectivity**	The degree to which the emotions expressed in a given text document are based on or influenced by personal feelings, tastes, or opinions.	(Pang and Lee, 2006; Cambria *et al.*, 2013; Lui, 2015)
6	**Opinion: Sentiment - Polarity**	The overall orientation (Positive, Negative or Neutral) expressed in a given text document.	(Pang and Lee, 2006; Cambria *et al.*, 2013; Lui, 2015)
7	**Relatedness: Cyber Relatedness**	The degree to which a given text is related to a cyber event.	Author
8	**Relatedness: Flu Relatedness**	The degree to which a given text document is related to a flu event as expressed in the experimental scenario.	Author
9	**Relatedness: Fire or Accident Relatedness**	The degree to which a given text document is related to a fire event as expressed in the experimental scenario.	Author
10	**Entropy**	The average degree of uncertainty associated with a given text document.	(C E Shannon, 1948; Bentz, 2016)
	Channel Congestion	The number of unique cyber personas engaged in a conversation within a given time window T.	(Cleveland and Sun, 1995)
	Channel Traffic	The number of messages flowing through a channel within a given time window T.	(Cleveland and Sun, 1995)
	Emotion Detection	The degree of emotions (one of Joy, Sadness, Anger, Disgust, Trust, Anticipation) expressed in a given text document	(Norcross, Guadagnoli and Prochaska, 1984)

Table 5-8: Features from the Social Dimension

In conclusion, datasets were collected, transformed and aggregated to suit the demands of the experiment detailed in the scenario development.

5.3 ANALYTICAL FRAMEWORK

This section explains how the data collected in this research is analysed. It outlines a procedural process for analysing the data at each stage of the experiment. Each dataset at each stage of the experiment is passed through a series of tests (Stationarity, Gaussian and Co-integration Tests). This

section also goes further to explain the criteria for feature selection, vector autoregressive model order selection and the final model selection. Additionally, techniques for model validation, model prediction and error propagation to the next phase of the kill-chain are also critically discussed. Following the experimental design, the analysis is designed to address the hypothesis developed in the theory and experiment. The following steps outlined in this framework are undertaken for each phase of the experimental analysis.

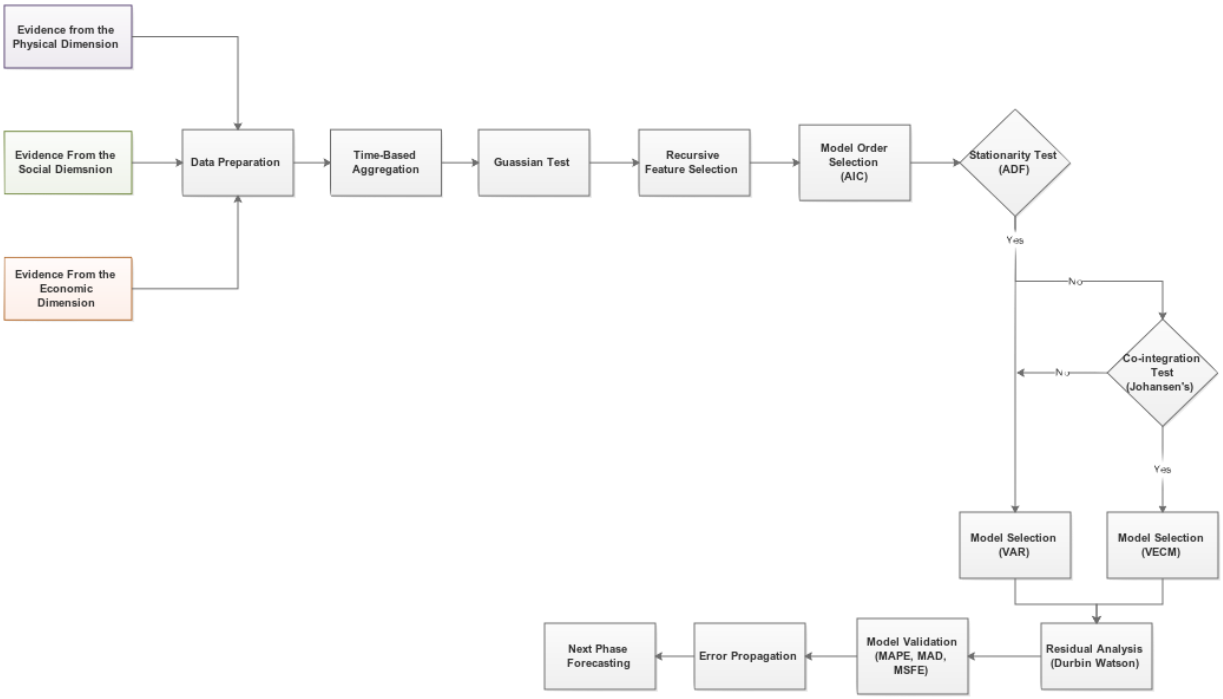

Figure 5-2: Experiment Analytical Framework

The methods depicted in above analytical framework and how they are used in the experimental design is further discussed in section 5.3, sub-sections 5.3.1 – 5.3.11.

5.3.1 Stationarity Test

A stationarity test was carried out for each variable on each dataspace to ensure data did not violate pre-assumptions of stationarity within vector autoregressive models (Luetkepohl, 2011). Here, we test for trend stationarity and unit root processes to determine the proper transformation technique to be applied to each feature. The Augmented Dickey-Fuller's (Cheung and La, 1995) test for unit root processes was conducted on each variable, and the results are documented in the experiment results. The Augmented Dickey fuller's difference stationarity test conducted tests the null hypothesis for the presence of a unit root in each variable. Given a stochastic feature X, the augmented Dickey fuller tests provided a test hypothesis:

$$H_0: \text{X contains a unit root}$$

$$H_1: \text{X does not contain a unit root}$$

Unit root processes are always non-stationary integrated order of D where D > 0, but may not always have a trend. The trend stationarity tests the null hypothesis that a process is a sum of a long-run linear trend and an invertible noise sequence (Hamilton, 1994). The researcher uses a difference transformation for unit root processes and logarithm transformation for trend stationarity data.

On the network layer, all variables excluding the 'The Number of Source Ips', 'The Number of Destination Ips', 'Mean time to live', 'Mean Captured Bytes' and 'The number of IDS Warning' were stationary. All variables on the social dimension also return a significant p-value for null hypothesis

stationarity. None of the features on the economic dimension were stationary. Consequently, data transformation was applied to detrend and normalize features that did not conform to pre-assumptions of stationarity and normality under necessary conditions in the experiment.

5.3.2 Gaussian Test

The Gaussian test is also known as a test for normality. Additionally, a Gaussian test was completed on each variable to ensure that pre-assumptions of normality are not violated. This research uses Jarque-Bera's Skewness-Kurtosis test to ensure that all constructs were within the acceptable limit of skewness-kurtosis ranges. This test is chosen due to its robustness as compared to other generalist tests such as Anderson-Darling test and the Shapiro-Wilk test. The skewness-kurtosis test draws a contrast between the study data distributions and normal distributions (Hair, 2014). Assuming a bell-curve is plotted for the distribution of a random variable, skewness and kurtosis are known measures that go hand-in-hand. While skewness displays the direction to which the curve has shifted, kurtosis measures the robustness of the bell-curve. On one hand, measuring skewness delivers some insights into the asymmetry and balance of probability distribution of the data around its mean. For instance, a positively skewed distribution produces a bell-curve significantly shifted to the right. On the other hand, measuring the kurtosis delivers insights into the peakedness or flatness of the distribution. Therefore, positive kurtosis values suggest a peaked distribution while negative kurtosis values suggest a flatter distribution (Hair, 2014). Skewness-kurtosis critical values have been discussed and examined by many different academics (Massey, 1951; Cheung and La, 1995; Ghasemi and Zahediasl, 2012) and the general consensus for an acceptable range is between ±2.58 at a 0.01 significance level.

The researcher uses the one-sample Kolmogorov-Smirnov normality test to access both the levels of skewness and peakedness of each time vector in the model at each stage of the experiment. The Kolmogorov-Smirnov normality test is a nonparametric test for the similarity of one-dimensional probability distributions. The Kolmogorov-Smirnov normality test tests the null hypothesis that the cumulative distribution function of the sample is equal to the cumulative distribution function of an hypothesized ideal normal distribution. It quantifies the difference or distance between the observations of the observed sample and the observations of an ideal normal distribution. The test is conducted at a 95% confidence level with a two-tailed alpha level of 0.025. If the estimated p-value is less than the chosen alpha level, the null hypothesis that the sample is similar to a normal distribution is rejected. However, if the p-value is less than the chosen alpha level, the null hypothesis cannot be rejected. Therefore, significant p-values corresponds with non-normality of data.

Logarithmic transformations are applied where necessary to meet assumptions of normality. Only 2.4% of features on the network layer fell within the +2.58 and – 2.58 recommended range for both skewness and kurtosis tests. Most of the features on the social dimension fall within the acceptable range of skewness but are however seen to have a leptokurtic distribution. 83.3% of data on the economic dimension pass the tests for normality on both skewness and kurtosis.

5.3.3 Outliers

Outliers are described as being the most extreme points of data that have moderate to significant impact the effectiveness of model constructs and thus overall findings and conclusions. Outliers in time series data are often regarded as unexpected interventions or innovations from which various

types of outlying observations can be produced. Identifying outliers in time series data is not straightforward as the presence of outliers may lead to multiple assumptions in the data. The easiest way to identify outliers is a simple line plot of the differences of data points against time. Outliers are seen as massive drifts from the mean and are usually interpreted as 'anomalies' in the data. While conventional methods of multivariate data analysis suggest the removal of outliers (Hair, 2014), the presence of outliers in time series explain the pattern of drifts from the series mean (Chen and Liu, 1993). (Chen and Liu, 1993) develops a procedure for automatically detecting four types of outliers in time series data.

- ***Additive outliers (AO)*** appear as a surprisingly large or small value occurring for a single observation in time. Subsequent observations remain unaffected by any drift of an additive outlier.
- Similarly, to additive outliers, ***level shift (LS)*** outliers appear as surprisingly large or small values in the series however, all observations appearing immediately after the outlier move to a new level. This shift may affect subsequent observations and the effects may be permanent.
- ***Innovational outliers (IO)*** are characterized by an initial impact at a single point in time but with effects that linger over subsequent observations of the series. The effects of this outlier may increase over time.
- ***Temporal or Transient Change outliers (TO)*** are very similar to level shift outliers, however the effects of the outlier on the series diminishes exponentially over time with the series eventually returning to its normal levels.

5.3.4 Missing Data

Several approaches have been described by researchers for addressing the issue of missing data in time series data. Firstly, missing data in time series refer to points in time for which observations for the variable was not made or recorded. The primary approach for dealing with missing data is linked with understanding the pattern of missing data. This involves the researcher establishing the sources of the missing data in line with both random and non-random occurrences. Observation bias or intentional omission is said to be absent in the data if missing data is randomly distributed throughout time. The data used in this research can be analyzed for missing data on each dimension. On the physical dimension, merging the ids logs, network traffic logs and firewall logs led to sometime observations with missing data. Only time stamps with corresponding date time values across all logs were included in the final dataset. Similarly, on the economic dimension, days where stock prices were not observed were completely taken out of the data.

Another commonly used approach for dealing with missing data especially in continuous datasets is the multiple imputation model (Acock, 2012). The idea behind the imputation model is to extract the dynamics of relevant information from the observed portions of the data, to impute multiple values for each missing data item. Most methods compute and construct multiple datasets depending on the estimated uncertainty in imputing each value. Multiple imputation has been proven to work properly for datasets with at most 30-40 features.

5.3.5 Correlation Analysis

To examine the linearity of relationships between observed features, this study uses the bivariate correlation matrix at a 0.01 significance level (2-tailed). As shown in the diagrams below, the is a 53.4% significant correlation level (Pearson's correlation coefficient is between +0.7 and -0.7)

between features on the economic dimension. On the social dimension, only 10% significant correlation levels were observed while the physical dimension returned 13.4% significant correlation rates.

5.3.5.1 Intra-Dimensional Correlation

This inter-dimensional co-integration between features within each dimension may lead to multi-collinearity in future analysis. However, this problem is addressed in the feature selection stage.

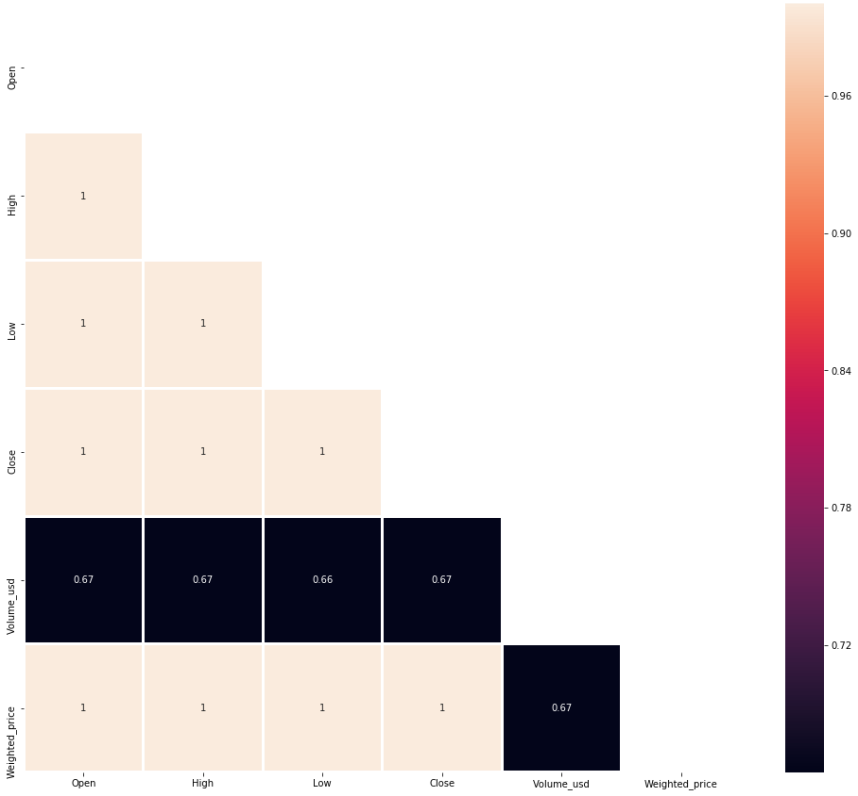

Figure 5-3: Inter-Dimensional Correlation on the Economic Dimension

Of the 7 (M=7) features on the economic dimension, 3 (Open, Close and High) are seen to be highly positively correlated while 'Volume' is seen to have just enough significance to be positively correlated with the former 3 features.

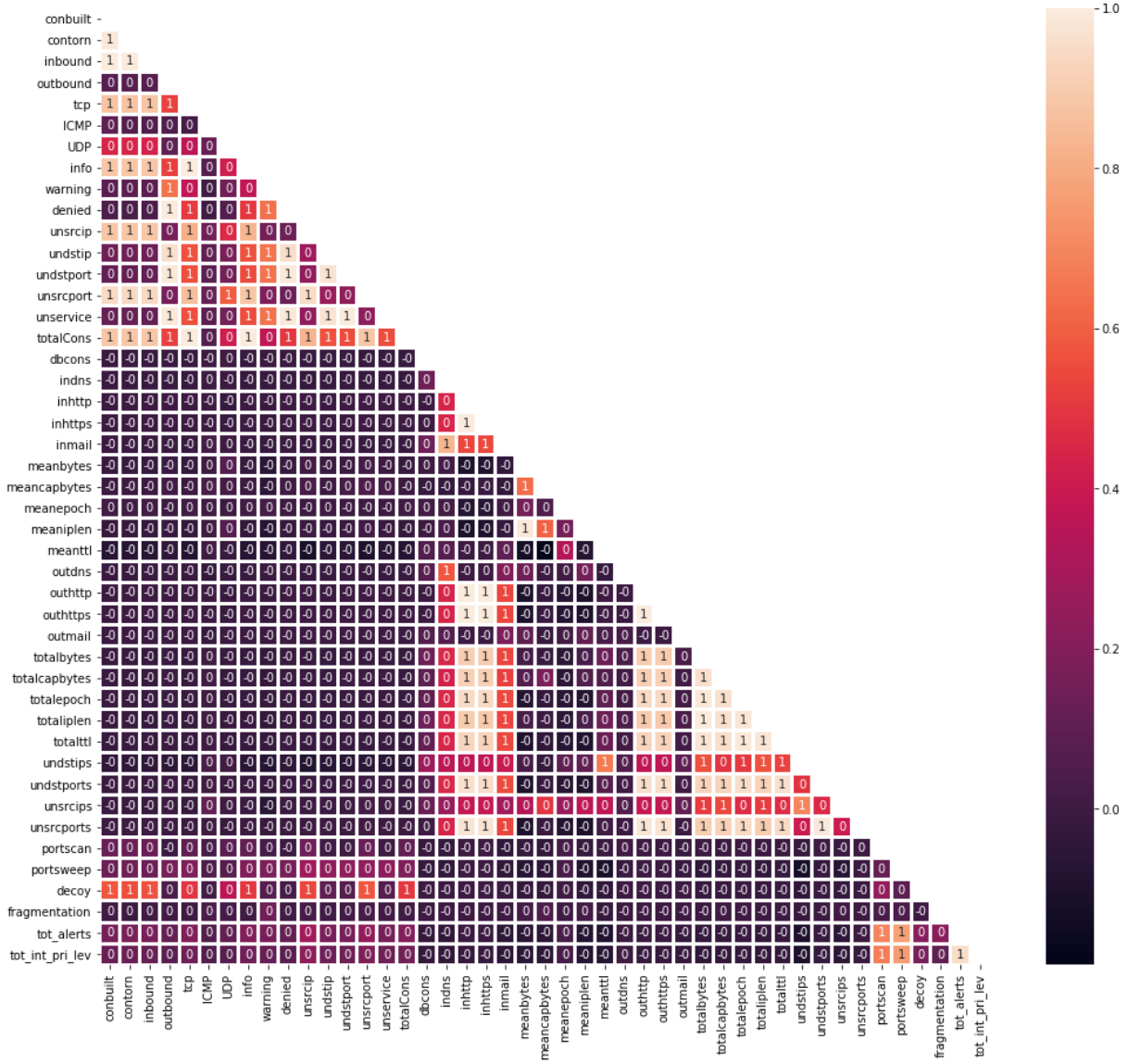

Figure 5-4: Inter-Dimensional Correlation on the Physical Dimension

The physical dimension consists of 45 (M=45) features. As expected, variables like 'The Number of TCP connections', 'The number of ICMP connections ('ICMP')', 'The number of UDP connections ('UDP')' are seen to have significantly high positive correlation coefficients with 'The total number of connections ('totalCons')', 'Number of Connection Built ('conbuilt') and torn down ('corntorn')' and 'The number of incoming ('inbound') and outgoing ('outbound') connections. Similarly, there is a significant linear relationship between the number of alerts from Intrusion Detection System and the number of applications running on the network.

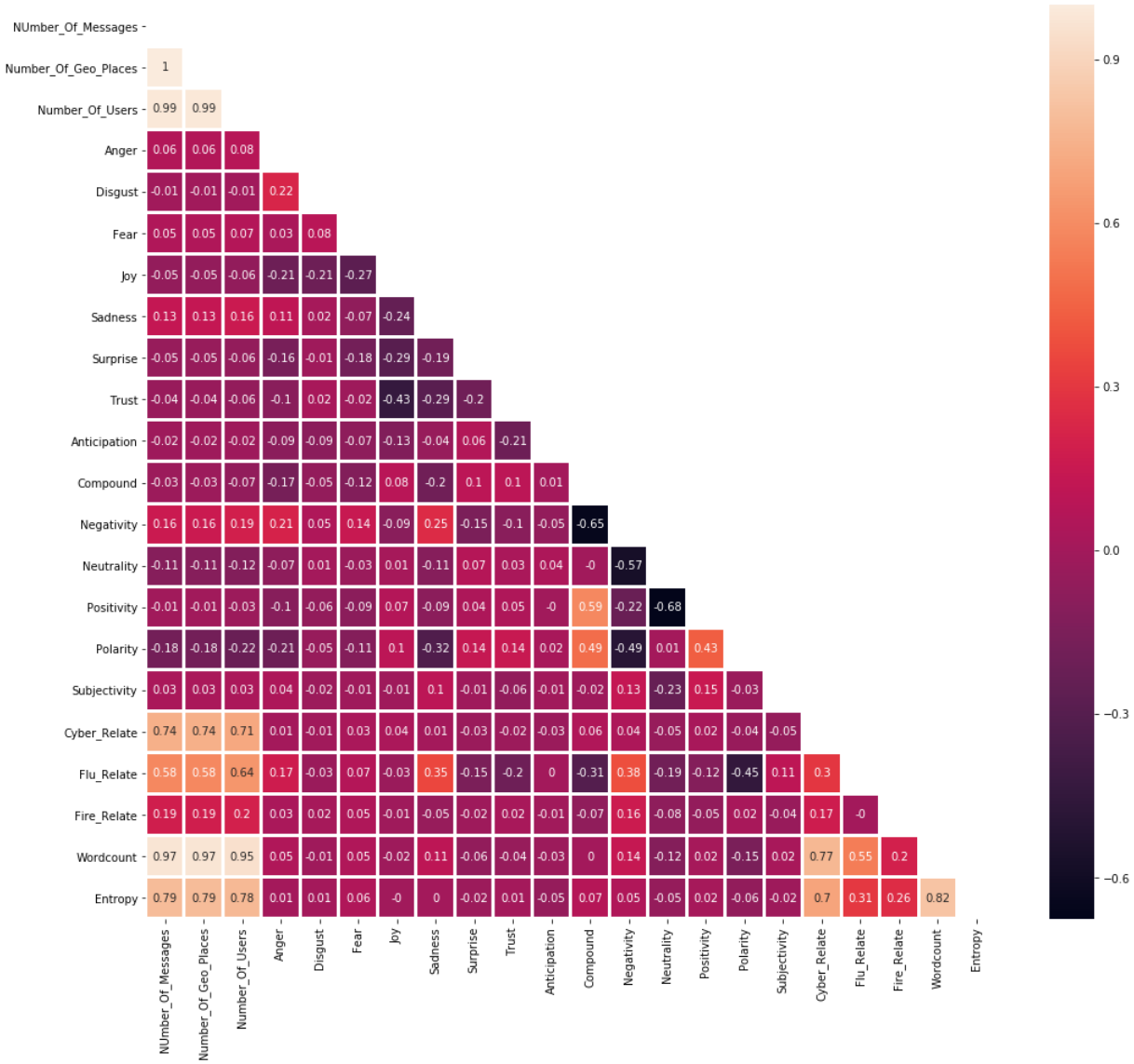

Figure 5-5: Inter-Dimensional Correlation on the Social Dimension

Of the 22 features on the social dimension, 3 sets of 3 variables are seen to be highly correlated. The 'Entropy', 'Chat Traffic' ('Number _Of_Messages') and 'Cyber Relatedness' ('Cyber_Relate') of variables are seen to all have significant correlation coefficients with each other. Similarly, the 'Chatroom Traffic' ('Number _Of_Messages') and 'Chatroom Congestion' ('Number _Of_Users') are also seen to have significantly high correlation coefficients (Pearson's Correlation Coefficient ≥ 0.7) with each other. These correlations are expected given that the 'quantification' techniques of microblogging messages in section 5.2.2.2.1 are function variations of the number of words in each text. Interestingly, there is an observed negative correlation between the happiness of the population and their trust in context of the scenario events (Event 1-section 4.5.3.3). Subsequently, the researcher demonstrates intra-dimensional correlation between features on varios dimensions of the information space.

5.3.6 Feature Selection

A fundamental problem of predictive analytics is dimension reduction (Hall and Smith, 1998) which is aimed at reducing the number of features under consideration to obtain a subset of principal variables. Feature selection is done to simplify models by making interpretation easier, reducing training time and computational cost by reducing the hypothesis search space, avoiding the curse of dimensionality (Verleysen and François, 2005) and to avoid overfitting and enhance model generalization. There are three main strategies for feature selection: filter methods, wrapper methods and embedded methods.

Wrapper methods build a predictive model for combination of feature subset. Wrapper methods include forward selection, backward selection and recursive selection. Additionally, wrapper methods are said to be very computationally intensive (a function of n X m) as they have to construct a model for every combination of subsets. Forward selection (Mehmood *et al.*, 2012) is an iterative model which starts with an original model with m=0 features. On each iteration, each feature that is not already in the model is tested for plausible inclusion in the model. Only features that best improves the performance of the model in the midst of other features already in the model are added to the candidate model. Model performance is set based on a pre-set threshold for the p-value of each feature added to the model, usually 0.05 i.e. a 95% significance level. One of the major drawbacks of forward selection is that significance of a variable is usually measured relative to other variables in the model. Therefore, each addition of a new variable, may render one or more variables already included in model as "non-significant". As opposed to forward selection, backward selection (Hall and Smith, 1998) starts with all features initially in the model. On each iteration, the least significant features of all features at a chosen critical threshold is dropped. Each iteration successively refits reduced forms of the model by applying the same rule until only significant variables are left in the model. Recursive Feature Elimination or Stepwise selection on the other hand allows the selection algorithm to move in both ways backwards or forwards. The algorithm drops and adds features at each iteration by keeping the best and worst of every model constructed. The process alternates between choosing the least significant feature to drop and re-considering all dropped features for re-introduction into the model. Therefore, this means that on each iteration, two separate hypotheses are tested, and two separate significance level are chosen for elimination and re-introduction.

Filter methods are usually used as a pre-processing step with feature selection independent of any learning algorithm. Instead, features are scored based on their performance in various statistical tests with the outcome variable. Filter methods are often deemed to be computationally less-intensive while still capturing the relative usefulness of each feature in the model. A correlation-based feature selection procedure uses a search algorithm to evaluate the merit of each subset of features. The correlation-based feature selection procedure measures the goodness of feature subsets with the usefulness of each individual feature in predicting the outcome class and also the inter-correlation among features. It eliminates one or more of inter-correlated variables as they are seen to be redundant and add no more useful information (not already known) to the model. Therefore, the goal is to attain a feature subset where each feature is highly correlated with the outcome variable but non-correlated with each other. (Hair, 2014) formalizes this measure of predictability VS redundancy as:

$$G_s = \frac{kr_{ci}}{\sqrt{k + k(k-1)r_{ii}}}$$

Equation 5-4: Model Feature predictability VS redundancy

Where kr_{ci} is a measure of predictability or usefulness, $\sqrt{k + k(k-1)r_{ii}}$ is a measure of redundancy, k is the number of features in the chosen subset and r_{ii} is the average feature inter-correlation and r_{ci} is the mean feature correlation of each feature with the outcome feature. Mutual information quantifies the amount of information gained about one feature upon the addition of another feature into the model. Therefore, it measures the amount of mutual dependence between two features. This method relies on the efficient estimation of the mutual dependence between a set of features and the outcome variable (Bonev, Escolano and Cazorla, 2008). Feature selection using mutual information is based on selecting only variables that share a mutual dependence with the outcome variables. This methodology measures the amount of information gained or lost by the addition or subtraction of each feature in the subset. Finally, a co-integration-based feature selection procedure also uses a search algorithm to evaluate the merit of each subset of features. The co-integration-based feature selection procedure measures the goodness of feature subsets with the probability that a linear combination of each feature with the feature set produces a stationary residual. It selects only variables that are co-integrated and would therefore improve the overall prediction performance of the outcome variable. While correlation is not a pre-cursor for co-integration, one drawback is that features may be correlated and co-integrated with the outcome variable leading to including redundant variables in the model. Finally, criterion-based procedures are also sometimes used as a feature selection technique. By selecting feature subset that maximizes any one of the information criteria outlined in the literature review, maximizes the adjusted R or minimizes the predicted sum of square errors.

This study combines both filter and criterion-based methods at each stage of the experiment to determine optimal feature set for predicting the selected outcome feature.

5.3.7 Model Order Selection

After only significant features have been selected for our candidate model, the next step is determining the correct order or lag length of the VAR(p) model. Parametric approaches to time series analysis require the need to estimate a model order. Determining the model order is an important step in vector autoregressive modelling and involves selecting a suitable lag length that optimizes the performance of the model. Model order selection by using statistical order-selection criterion was first introduced by Akaike (Akaike, Clements and Hendry, 1969) in fitting autoregressive models for prediction. The importance of correct estimation of model order is demonstrated by (Braun and Mittnik, 1993; Luetkepohl, 2011) who show that results of a VAR model whose model order differ from the true lag length are inconsistent. Various methods for appropriate model order estimation has been suggested over the years (Ventzislav and Lutz, 2005) resulting a number of lag order criteria selection methods: AIC, HQ, FPE, SC. (Ventzislav and Lutz, 2005) compare these various model order selection criteria commo0nly used in applied time series analysis and conclude that the Akaike Information Criteria (AIC) (Akaike, Clements and Hendry, 1969) produces the most accurate estimates for most VARs. However, the Hanan-Quinn criterion (HQC) (Hannan and Quinn, 2010) seems to be the most accurate for quarterly data and the Schwarz Information Criteria (Schwarz, 1978) seems to be the most accurate for small datasets (N <= 120). This study estimates a VAR order by simple OLS per equation and selects model order that maximizes model performance based on Akaike Information Criteria (AIC).

Most often in VAR lag length selection, symmetric lag lengths are used i.e the same lag length is used for every variable in the model. (Ozcicek and Douglas McMillin, 1999) suggested estimating VARs using asymmetric lags where the lag length of each variable may differ. The researcher also explores the possibility of asymmetric lag length selection.

5.3.7.1 ACF

The auto correlation function measures the serial correlation of a time variable with itself at two different points in time. It is the estimated similarity between observations of a time series as a function of the time lag between them. Formally, the ACF is a function of the covariance and standard deviation between a time series and past values of itself. The ACF of a time series x_t is given as:

$$\frac{Covariance\ (x_t, x_{t-h})}{Std.Dev\ (x_t) Std.Dev(x_{t-h})} = \frac{Covariance\ (x_t, x_{t-h})}{Variance\ (x_t)}$$

Equation 5-5: Auto-Correlation Function (II)

5.3.7.2 PACF

In general, the Partial auto correlation function (PACF) is a conditional correlation between two variables under the assumption that the values of some other feature in the feature set is taken into consideration as well. Formally, the Partial auto-correlation function is a function of the conditional covariance and variance between the variables of a regression model. It is given as:

$$\frac{Covariance\ (y, x_{3\ |\ x_1, x_2})}{\sqrt{Variance}\ (y\ |x_1, x_2)\ Variance\ (x_3\ |x_1, x_2)}$$

Equation 5-6: Partial Auto-Correlation Function (II)

The formula above represents the 1st order partial auto correlation i.e. the conditional correlation between values of the series one time period apart conditional on the knowledge of the values in between. If the series is stationary, the variances $\sqrt{Variance}\ (y\ |x_1, x_2)$ and $Variance\ (x_3\ |x_1, x_2)$ should equal each other. The partial auto correlation plot indicates the strength of the relationship between different lags of values of a time series. The partial auto correlation plot can also be used for visual inspection of appropriate model order as it shows points at which series are relevant in predicting its current values.

Finally, the auto-correlation and Partial auto-correlation functions are useful in to determine the appropriate lag orders for time series models. Consequently, this addresses the 'sensitivity to useful time frames' limitation of the granger causality test. By using the appropriate lag order, the definite lags at which co-integration occurs between the two-time series identified.

5.3.8 Co-integration Testing

After appropriate features set has been duly selected, the reduced features set is tested for co-integration. Given the features set Ω, the researcher estimates the co-integrating rank *r* of a multivariate time series data. At each stage of the experiment, the researcher constructs a co-integration matrix of all selected features and derives the co-integrating rank *r* of the feature set. The co-integration procedure measures the goodness of each feature combination in the feature subsets with the probability that a linear combination of one or more features including the outcome feature produces a stationary residual. The co-integrating rank *r* is a measure of how many unique combinations of variables are co-integrated.

In practice, there are two main approaches to co-integration testing the Engle and Granger approach (Engle *et al.*, 2017) and the Johansen's approach (Johansen, 2000) which correspond to bivariate and multivariate co-integration testing in time series analysis. The researcher follows the following steps to test for co-integration in the final feature set:

- The researcher conducts an augmented dickey fuller test on each variable in the model with the following hypothesis:

  H_0: The process contains a unit root.
  H_1: The process contains no unit roots.

  The researcher eliminates variables where the null hypothesis is not rejected as co-integration implies a common stochastic trend between two variables.
- The researcher performs a Johansen's co-integration test on the selected feature set to determine the co-integrating rank *r* of the feature set. The Johansen's test can be conducted either using the trace test or the eigen test. Both tests are similar, however differ in the specification of the null hypothesis. The null hypothesis for the trace test is that the number of co-integrating vectors is less than the number of variables while the eigen tests the null hypothesis that the co-integrating rank is exactly equal to the number of variables. The test is conducted sequentially for r=1,2,3..N and the first non-rejection of the null hypothesis is taken and the co-integrating rank r.

- To determine which variables are co-integrated, the researcher constructs a linear regression model for each combination of the untransformed features in the feature set using the Engle & Granger approach (Robert F. Engle and Kenneth F. Kroner, 1995).
- The researcher conducts an augmented dickey fuller's stationarity test on the residuals of the resulting OLS regression model with a null hypothesis and alternative hypothesis as stated above.

5.3.9 Model Selection

The Engle & Granger approach suggests a two-step approach to co-integration testing: testing for co-integration and if variables are co-integrated, estimate the error correction model to investigate the long-run relationship. After a test for co-integration has been done the I(1) features in the selected feature set, a framework for estimating the model is selected. This selection is based on the results of the preceding co-integration test.
The VAR model (**2.9.6**) is used when there is no co-integration between any combination of variables in the system i.e $r=0$. If all variables are stationary, the VAR is fitted in the levels to the series otherwise, the VAR is then estimated using the stationary transformed I(0) features.
However, if the variables are I (1) and co-integrated ($r > 0$), then the system of equations is modified to allow for co-integrating relationships among the features. In traducing the co-integrating relationship leads to the estimation of a Vector Error Correction Model (Section **2.9.7**).

5.3.10 Residual Analysis

Residual analysis provides a general approach to access the quality of models by checking if the model has achieved its goal of explaining as much variation in the dependent variable as possible. Additionally, in response to the limitations of the assumption of linearity of Granger Causility testing as outlined in section 4.4.1, testing the normality of residuals is a good way to ensure that the relationship between the time series under observations meet the assumption of linearity. Residuals are left-overs of the model after the variation in dependent variable using the independent variable has

been explained. Ideally, residuals should be stationary, unstructured and as close to zero as possible i.e residuals for each forecast must be as close to the actual value as possible. To this end, residuals are expected to be stationary with mean 0 and equal variance over time. To this effect, the researcher tests the residuals for each parameter of the model for stationarity, normality and serial auto-correlation. The researcher employs two methods for residual analysis, visual methods and parametric testing methods. Visually, the researcher uses the Q-Q plot and the cumulative frequency plot to visualize the distribution of residuals. Additionally, the researcher uses auto-correlation and partial auto-correlation plots to test for serial auto-correlation in the residuals. Using the Shapiro-Wilk's test, the Augmented-Dickey Fuller test and Durbin-Watson test, the researcher tests for normality of residuals, stationarity of residuals and serial auto-correlation of residuals.

The Durbin-Watson test, otherwise known as the DW statistic is a test for autocorrelation in the residuals from a statistical model. However, the test can be extended to test for autocorrelation in the series itself. The DW test estimate usually falls between 0 and 4, 0 indicating a highly positive autocorrelation and 4 indicating a highly negative autocorrelation. A value of 2 is a point of neutrality that indicates no autocorrelation in the series under observation. The Durbin-Watson test is estimated as:

$$DW = \frac{SDES}{SSE}$$

Equation 5-7: Durbin Watson Test For Autocorrelation

Where

$$SSE = Sum\ Of\ Squared\ Erros = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \text{ and}$$
$$SDES = Sum\ of\ Squared\ Error\ Differences = \sum_{i=1}^{N}\Delta SSE^2.$$

5.3.11 Model Validation

In time series analysis, model validation substantiates that a computerized model within its domain of applicability possess a satisfactory range of consistency within the intended application of the model. Model validation involves testing the performance of a model across multiple samples of forecast values using new data. At each stage of the experiment, the researcher estimates a value for each of the following performance measurements:

5.3.11.1 Goodness of Fit- R^2 and Adjusted R^2:

Traditionally, the linear regression model calculates an equation that minimizes the distance between the equation line and all data points in the model. R^2, also known as the co-efficient of determination, estimates the how close the data points are to the regression line. The R^2, is simply interpreted as the percentage of the variation in the response variable explained by the model. The estimates for R^2 usually fall within the ranges of 0%, indicating that the model explains none of the variation in the response variable and 100%, indicating that the model explains all the variation in the response variable. Higher R^2 corresponds to better fits. However, it is also important to check for spurious regressions in model by comparing R^2 estimates to the Durbin-Watson statistic. Given that the Vector Autoregressive model is an OLS per equation model, the average R^2 for the VAR(p) model construct at each stage is the average of the individual R^2 for each parameter equation in the model. Unlike the R^2, the adjusted R^2 is a modified version of the R^2 that considers the number of explanatory variables in the model. The adjusted R^2 only increases if a new explanatory variable added to the model improves the model more than would be expected by chance and decreases otherwise. One major disadvantage of measuring model performance using the R^2 estimate is that it assumes that all explanatory variables explains some variation in the dependent variable. It therefore computes the

percentage of explained variation like all independent variables actually affects the dependent variable. On the other hand, the adjusted R^2 estimates the percentage of variation explained by only those independent variables that are actually relevant for predicting the dependent variable.

5.3.11.2 Evaluating Forecast Accuracy

Evaluating the accuracy of forecasts generated by a model is an important step in structural time series analysis. Forecasts are the models predicted values, and forecasts are evaluated by accessing how close to the actual observations they are. Evaluating forecasts accuracy is usually done in two steps: in sample forecasting and out of sample forecasting. In-sample performance evaluation is done by comparing the model's forecast to the values of the actual data. Out of sample performance is conducted to evaluate the model's performance on new data i.e data not used in model construction. At the start of any time series analysis, a number of n last observations is taken out of the time series and a prediction model is built with n-ahead forecasts. The n-ahead forecasts values are then compared to the reserved last n-observations to evaluate models out of sample performance. Rob Hyndman (Hyndman and Koehler, 2006) pushes this further by defining a procedure for cross-validating time series based on a rolling forecast of every n-ahead model predictions. Each n-ahead data observation acts as a new test observation. The overall forecast accuracy is then estimated by computing the average over each n-ahead prediction.
Measuring the accuracy models is typically done by measuring the size of the errors in the predictions generated by the model. Some model performance measurement techniques based on error analysis are given below.

5.3.11.2.1 SSE (Sum Of Squared Errors)

The sum of squared errors also known as the residual sum of squares or the sum of squared residuals is the total addition of the squares of errors (difference between the predicted value and the actual values). Theoretically, the sum of squared errors measures the discrepancy or similarity between the data and the estimated model values. Formally, the sum of squared errors is given by:

$$SSE = \sum_{i=1}^{N} (Y_i - \hat{Y}_i)$$

Equation 5-8: Model Sum Of Squared Errors

Where N is the number of observations, Y_i is the value of the ith data observation and $\hat{Y}$ is the value of the ith predicted value. In a Vector Autoregressive model, a simple OLS equation is built for each parameter in the model. This research estimates the SSE for each parameter in model and averages the SSE across all parameters to access the total SSE of the model. The total SSE is estimated and reported both in sample and out of sample.

5.3.11.2.2 MAE (Mean Absolute Error)

Since errors may be positive or negative, the mean absolute error is a measure of the size of errors irrespective of the direction of the errors. The mean absolute error is therefore the average of the absolute values of errors $|\varepsilon_i| = |y_i - x_i|$ where y_i are predicted values and x_i are the actual values of the series. The MAE is a scale-dependent measure of accuracy (A function *f*(.) is scale-invariant if it yields the same result for argument *x* as it does for argument *cx*, where *c* is some positive constant).

The MAE is a common measure of accuracy in time series analysis sometimes referred to as the 'Mean Absolute Deviation'. Formally, the Mean Absolute Error is given as:

$$MAE = MEAN(|\varepsilon_i|)$$

Equation 5-9: Model Mean Absolute Error

Where ε_i is the error term associated with the ith prediction and the ith value of the series. This research estimates the MAE for each parameter in model and averages the SSE across all parameters to access the total MAE of the model. The total MAE is estimated and reported both in sample and out of sample.

5.3.11.2.3 RMSE (Root Mean Squared Error)

The root mean square error also sometimes called the root mean squared deviation measures the deviations of the predicted values format eh actual values. The value is estimated by finding the square root the mean of the squared error. The RMSE represents the standard deviation of the residuals or predicted errors for out-of-sample computations. The RMSE attempts to aggregate the size of prediction errors for each time observation into a single measure of predictive power. RMSE is scale-dependent (A function *f*(.) is scale-invariant if it yields the same result for argument *x* as it does for argument *cx*, where *c* is some positive constant) and is therefore used to compare errors of models on the same dataset. The TMSE is formally the square root of the average of the squared errors and will therefore always be a positive number. The effect of each error observation on the total RMSE is directly proportional to the size of the squared errors, consequently, RMSE measures are extremely sensitive to outliers. Formally, the RMSE is given as:

$$RMSE = \sqrt{MEAN(\varepsilon_i^2)}$$

Equation 5-10: Model Root Mean Squared Error

Where ε_i^2 is the square of the ith error for all *i-observations through N*. This research estimates the $RMSE$ for each parameter in model and averages the $RMSE$ across all parameters to access the total $RMSE$ of the model. The total $RMSE$ is estimated and reported both in sample and out of sample.

5.3.11.2.4 MAD (Mean Absolute Deviation)

The Mean Absolute Deviation is a scale-dependent evaluation technique that measures the size of prediction and forecast errors in units. The MAD measurement estimates the average of the absolute values of the forecast errors over the observed time period. This method is independent of the direction of the errors and therefore avoids positive and negative errors cancelling out each other. The Mean Absolute Deviation is defined formally as:

$$MAD = \frac{1}{N}\sum_{t=1}^{N} |\, F_t - A_t \,|$$

Equation 5-11: Model Mean Absolute Deviation

Where N is the number of observations, A_t are the actual values of the time series and F_t are the forecast values generated by the model. This research estimates the MAD for each parameter in model and averages the MAD across all parameters to access the total MAD of the model. The total MAD is estimated and reported both in sample and out of sample.

5.3.11.2.5 MAPE (Mean Absolute Percentage Error)

The MAPE (Mean Absolute Percent Error) measures the size of the error in percentage terms. It is calculated as the average of the percentage of error terms for each ith associations with the predicted and actual values. The MAPE is often used in practice as it is intuitive to interpret in terms of relative errors. The MAPE is a percentage error and therefore has the advantage of being scale-dependent and insensitive to extreme values. Percentage errors however have the drawback of being infinite or undefined if the actual observations are zero for any time period of interest. Percentage errors are also criticized for having extreme values when the actual values are closer to zero (Hyndman and Koehler, 2006). For real world applications, the MAPE is frequently used when the actual values of a series are way above zero. Formally, the Mean Absolute Percentage Error (MAPE) is the mean of the absolute values of the percentage errors estimated for each observation the data. It is given as:

$$MAPE = \frac{100\%}{N} \sum_{t=1}^{N} \left| \frac{A_t - F_t}{A_t} \right|$$

Equation 5-12: Model Mean Absolute Percentage Error

Where N is the number of observations, A_t are the actual values of the time series and F_t are the forecast values generated by the model. This research estimates the $MAPE$ for each parameter in model and averages the $MAPE$ across all parameters to access the total $MAPE$ of the model. The total $MAPE$ is estimated and reported both in sample and out of sample.

5.3.11.3 Co-integrating Ranks

In addition to the measures of accuracy and performance outlined above, this research also estimates values for inter-dimensional co-integrating ranks and intra-dimensional co-integrating ranks. These rank values are used to assess the level of integral dependence between parameters of the model. The level of co-integration between variables of the model estimates the predictive power of the structural model for each parameter in the model.

5.3.11.3.1 Intra-Dimensional Co-Integrating Rank

Given a multi-dimensional dataspace Ω with time series vectors representative of incidents in cyberspace across various dimensions of cyber space X_t, Y_t and Z_t, if the co-integrating rank r for any pairs of variables across X_t, Y_t and Z_t is greater than 1, we say these variables are intra-dimensionally co-integrated. $r_\Omega > 0$.

5.3.11.3.2 Inter-Dimensional Co-integration

Given a multi-dimensional dataspace Ω with time series vectors representative of incidents in cyberspace across various dimensions of cyber space X_t, Y_t and Z_t, if the co-integrating rank r for any pairs of variables within X_t, Y_t or Z_t is greater than 1, we say these variables are inter-dimensionally co-integrated. $r_X > 0 \mid r_Y > 0 \mid r_z > 0$.

5.3.11.4 Percentage Parameter Prediction Performance

Given that the VAR model is a set of structurally represented K OLS equations, where K = number of parameters in the model, each parameter is predicted per equation. This research measures the performance of the model for each parameter in the model.

SN	Analytical Step	Justification
1	Evidences from the social, physical and economic dimension	Integrates data that represents activities on each of these dimensions
2	Data preparation and Time-Based aggregation	Cleans data and integrates on a similar timeline
3	Normality Test	Test Data meets the assumption of normality. If not, data is student normalized.
4	Feature selection	Features relevant to predicting the target variable are selected.
5	Model Order Selection	The correct lag of the model is selected using the ACF testing.
6	Stationarity Test	Test data meets stationarity assumption. If not, data is integrated once.
7	Co-integration Testing (If Stationary)	Test for co-integrating rank between time series data.
8	Model Selection	Select appropriate model
9	Residual Analysis	Analyses the size and distribution of error to test the goodness of the model.
10	Model Validation	Tests how well the model performs in identifying entangled relationships.

Table 5-9: Summary of Analytical Framework

5.4 DATA ANALYSIS

5.4.1 Stage 1: The Antecedents Phase

Stage 1 of the experiment integrates evidences from the social and economic dimension to spot antecedents to cyber-incidents. The rationale for conducting this stage of the experiment is to establish entanglements between events on the social dimension and economic dimension of cyberspace. The researcher uses series from the economic dimension and the social dimension in constructing a monitoring model. The endogenous variable selected at this stage is the 'Weighted Price Index' of the company's share prices. Both datasets were integrated on a similar timeline with the aim of predicting events of the last month on the economic dimension. The merging of both datasets produced a single dataset with 26 features and 30239 observations. Additionally, the data was split into two removing the last 120 observations (the last hour) from each feature in the dataset. The out of sample reliability of the constructed model will be tested on observations from the last hour using a N-ahead prediction strategy where N=120. This is done to ensure reliability of out of sample prediction performance testing.

A simple intra-dimensional correlation analysis of the 26 features on both dimensions as shown in the figure below, reveals a slight correlation between features on both dimensions. For example, the average wind speed on the social dimension is seen to have a correlation coefficient with several features on the economic dimension. The black boxes in the figure below groups features into hierarchical clusters. The spread of the probability of a cyber-incident is seen to be clustered with the volume of share prices on the economic dimension. Figure 5-6 below shows the values for the intra-dimensional correlation between features on the economic and social dimension.

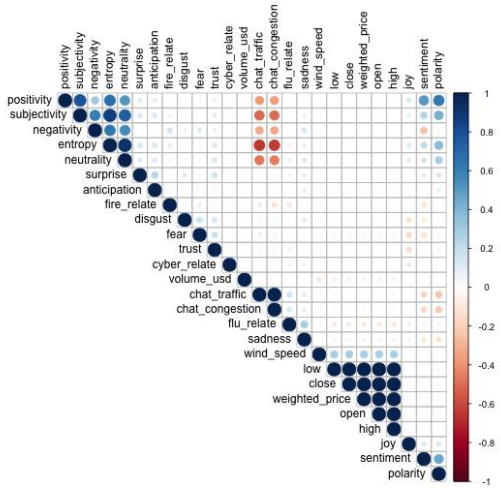

Figure 5-6: Intra-Dimensional Correlation Coefficients (Antecedents Phase)

The figure above shows the intra-dimensional correlation between features across the economic and social dimensions of the information space represented in this experiment. The correlation coefficients are on a scale of +1 to -1 as shown by the color bar. Highly positively correlated pairs of variables tend towards +1 while highly negatively correlated pairs of variables tend towards -1. Multiple pairs of variables are seen to be highly positively or negatively correlated. For example, the group of features, Cyber relatedness ('cyber_relate'), Entropy, chat room congestion ('chat_congestion') and chat room traffic ('chat_traffic') posted per minute are all seen to be highly positively correlated. Additionally, the group of features Weighted price ('weighted_price'), Open prices ('open'), closing prices ('close'), High ('high') and Low ('low') prices are also seen to be highly correlated. Consequently, one or more of pairs of features that are seen to be highly correlated with each other (where Pearson's correlation coefficient $\geq +0.65$ or ≤ -0.65) are removed from the dataset. The correlation elimination process reduces the feature set by removing one or more of a set of highly correlated variables (Hall and Smith, 1998). The assumption is that information provided by highly correlated variables can be provided to the model by a single one of those variables therefore reducing redundancy to achieve a parsimonious model. The results of the filtered feature set after a correlation elimination produces a smaller feature set of 15 features as shown in the figure below, eliminating 11 highly correlated features.

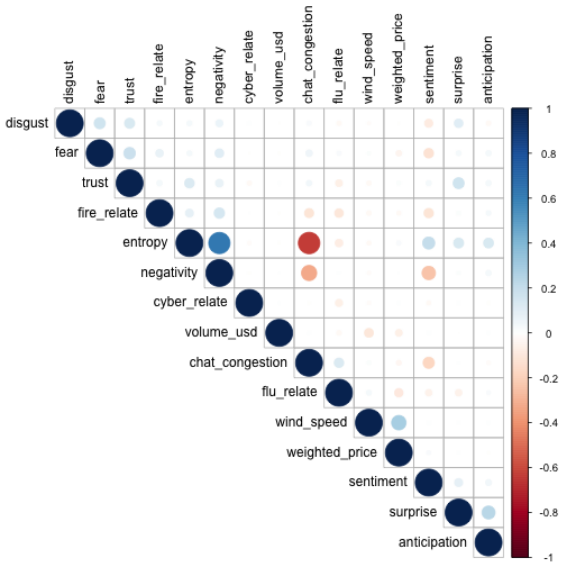

Figure 5-7: Correlation Co-efficient of Reduced Feature set (Antecedents Phase)

The figure above shows the intra-dimensional correlation between features across the three dimensions of the information space represented in this experiment for the reduced feature set. There are some correlations identified between features across the economic and the social dimension. For example, the little identified correlation between the 'wind_speed' and the 'weighted_price' of Delish corporation and also the 'wind_speed' and company's stock volume ('volume_usd') in dollar terms. Given the scenario-a epidemic possible spread by wind', these correlations are expected given information from other variables such as emotions from the population regarding the company and its effects on company's share prices'. This reduced set of 15 variables including the endogenous variable, 'weighted_price' are selected for further feature selection methods. To further reduce the number of features, the researcher applies a recursive features selection method to select only features that significantly reduce the error of prediction in the endogenous feature. The recursive feature selection stage further reduces the uncorrelated feature set to only those variables that minimizes the prediction error for the chosen endogenous variable 'weighted_price'. To begin the recursive feature selection from the uncorrelated feature set, the researcher starts by building an ordinary least squares regression linear model with the 'weighted_price' as the endogenous feature predicted by all other features in the reduced feature set above as shown in the figure above. The recursive feature selection searches for a model that optimizes the Akaike Information (AIC) and reduces errors. This step further reduces the number of features leaving 8 features. The resulting features are tabled below and used in further analysis. Lastly, the researcher creates generic methods for calculating the importance of each feature in predicting the outcome feature. This is used to examine the contribution of each feature in the dataset in predicting the outcome feature. This last step in the feature selection phase is elimination by variable importance (Gevrey, Dimopoulos and Lek, 2003). The absolute value of the t-statistic for each feature is used to compute the unique contribution of each input-feature to the model and a cut off of 1 is chosen (Noppamas, Seree and Kidakan, 2014). From the figure below, it is observed that the most important features for predicting the weighted price index are the average wind speed, flu relatedness of microblog messages and the chatroom congestion.

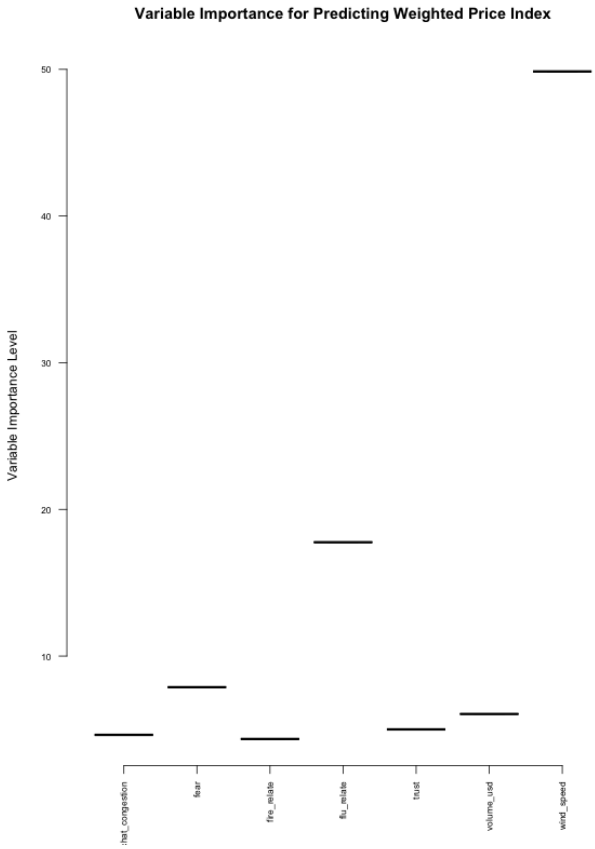

Figure 5-8: Variable Importance Of Features (Antecedents Phase)

The last phase of the feature selection, depicted in figure 5-8 above, shows that the 7 homogenous features selected by the AIC are extremely useful for predicting the outcome feature. The flu relatedness of microblogging feed 'flu_relate' is seen to be the most important variable for monitoring the antecedent phase of the experiment within the context of the scenario. Event-based features such as flu-relatedness ('fire_relate') and fire_relatedness ('fire_relate') are important for identifying antecedents to the proliferation of kill-chains.

The next phase of the analytical framework involves selecting appropriate orders for the intended model. This involves selecting an autoregressive order, testing for seasonality in each feature and selecting a moving average as discussed in the analytical framework.

The VAR model order selection phase seeks to select the appropriate significant lag at which features are serially correlated. The researcher uses an iterative solution with the maximum of $\frac{N}{2}$, where N is the number of observations in the feature set. At each iteration, the researcher builds a model for each equation in the VAR and measures the AIC. Finally, the researcher selects 36 lags (half-hour) as this is the lag length with the optimal AIC. The partial auto correlation plots below show this to be valid for each feature selected.

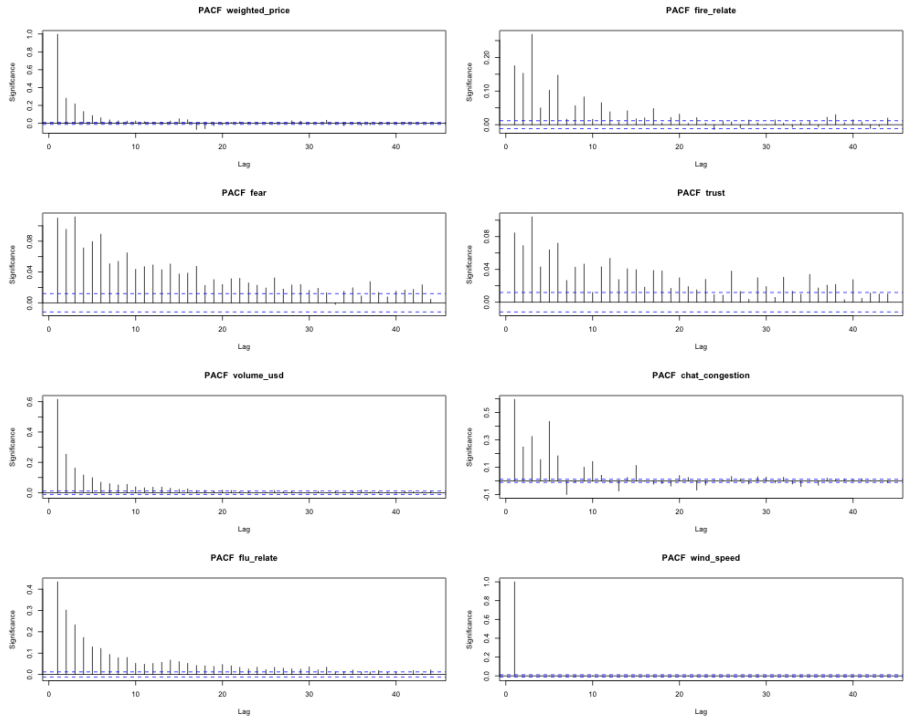

Figure 5-9: Partial Auto Correlation of Selected Features

In addition to selecting appropriate VAR lag order, the researcher also tests each time series in the feature set for seasonality. In order to achieve this, the researcher employs a time series decomposition technique. As discussed in the literature review, time series decomposition works by splitting a time series into its three main components: the trend, the seasonal movement and the error terms. Seasonal patterns are expected to repeat with a fixed period of time. An additive decomposition method was chosen as the magnitude of the series does not increase with the series. After applying the Fourier transform for detecting seasonality, four features exhibited seasonal trend. Emotion probabilities such as 'Fear', 'Trust' and the chatroom congestion has a daily seasonality with a time period of 1446 minutes. Additionally, the cyber relatedness of texts in microblogging feeds are seen to have a 12-hour seasonal time period.

A stationarity and normality test is conducted on the 9 features and results are tabled above. The results of the stationarity test show a significant p-value for 7 of the 9 features. These selected features are seen to have a constant mean and variance over the selected period of time. However, two features, the weighted price and wind speed are seen to be non-stationary. This feature is transformed by mean normalization as follows:

$$\frac{x - \mu}{\sqrt{\sigma}}$$

Where μ is the feature is mean and σ is the feature variance.

	Stationarity Test			Normality Test			
	P-Value	**Test Statistic**	**Stationary**	**Skewness**	**Kurtosis**	**Kolmogorov–Smirnov Test Statistic**	**K-S P-Value**
Weighted Price	0.09	-3.20	False	0.07	-0.93	1.00	0.00
Fire Relatedness	0.01	-17.34	True	2.05	7.86	0.50	0.00
Fear	0.01	-14.84	True	1.14	3.75	0.50	0.00

Trust	0.01	-17.93	True	0.63	1.59	0.50	0.00
Volume	0.01	-15.60	True	0.97	0.51	1.00	0.00
Chatroom Congestion	0.01	-12.11	True	2.15	7.87	1.00	0.00
Flu Relatedness	0.01	-8.42	True	2.10	9.80	0.51	0.00
Average Wind Speed	0.50	-2.18	False	0.16	0.20	1.00	0.00

Table 5-10: Features' Stationarity and Normality Test Results

The Johansen's co-integration test shows that there are 5 co-integrating relationships between features in the feature set. The Engle and Granger test for co-integration was used to determine co-integration between each pair of features in the feature set. To achieve this, the researcher constructs a linear regression model between each pair of features and conducts a stationarity test on the residuals of the resulting model. The Engle Granger test shows significant p-values for three features: the weighted price index, the chatroom congestion and the population fear derived from microblog feeds. The figure below shows a long-run equilibrium relationship between these three features with this relationship strongest towards the end of the period of observation.

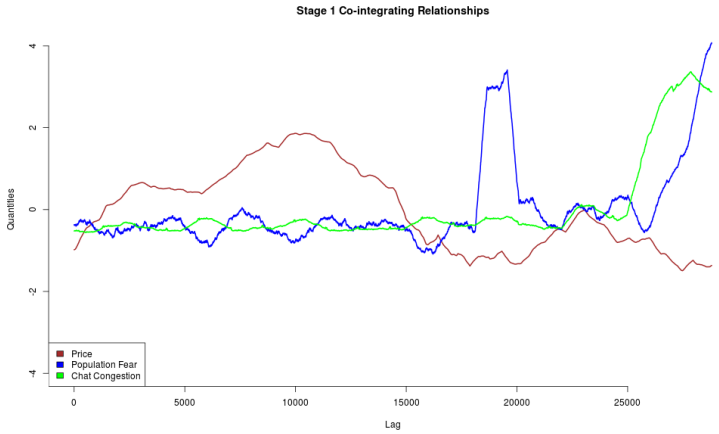

Figure 5-10: Co-Integrating Relationships Identified in the Antecedents Phase

Given the presence of multiple co-integrating relationships and some non-stationary vectors, a VAR model that takes into consideration these relationships are used. A Vector Error Correction Model (VECM) is fitted to the data in the levels. The VECM with order of 30 lags was estimated utilizing the OLS per equation in the model with 8 parameter OLS models where each feature is predicted by its on lags and lags of the other 7 features in the model. The density distribution plot below shows the residuals from the fitted VECM model. The error terms i.e. the residuals are expected to be independently distributed across each equation and serially uncorrelated. The residuals are experimental errors derived by finding the difference between the observed data points and the predicted data points. To ensure that these assumptions are met, the researcher tests the hypothesis of white noise residuals using a normality test on the residuals of the VECM model. The stationarity and normality test results for the residuals are shown below.

	Stationarity Test			**Normality Test**			
	P- Value	**Test Statistic**	**Stationary**	**Skewness**	**Kurtosis**	**Kolmogorov–Smirnov Test Statistic**	**K-S P-Value**

Weighted Price	0.01	-29.49	True	-0.63	131.30	0.44	0.00
Fire Relatedness	0.01	-30.53	True	1.74	8.06	0.10	0.00
Fear	0.01	-32.27	True	1.10	4.14	0.06	0.00
Trust	0.01	-30.77	True	0.67	1.95	0.04	0.00
Volume	0.01	-30.00	True	0.71	1.26	0.11	0.00
Chatroom Congestion	0.01	-31.38	True	0.54	4.66	0.15	0.00
Flu Relatedness	0.01	-34.40	True	0.31	2.57	0.08	0.00
Average Wind Speed	0.01	-29.56	True	-15.98	3208.72	0.50	0.00

Table 5-11: Residual Stationarity and Normality Test Results (Antecedents Phase)

The distribution of the residuals for the weighted prices and average wind speed as shown in the figure below indicate a leptokurtic distribution with a strongly narrowly peaked curve. This indicates that the model is able to capture most of the trend in the endogenous variable. The model performs moderately in capturing the trends in the other features as the residuals are seen to be slightly skewed to the left.

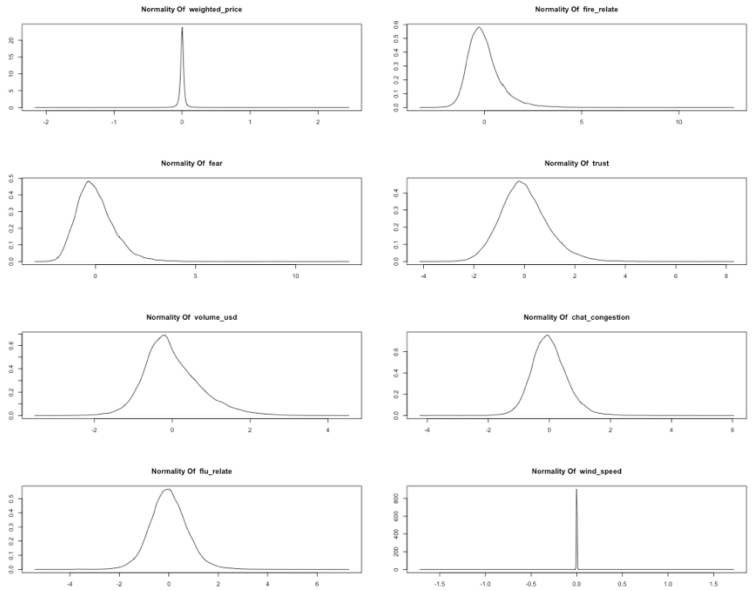

Figure 5-11: Density Distribution of Residuals (Antecedents Phase)

Lastly, after observing the density distribution of the residuals, the researcher derives a normal probability plot for each residual from the structural model. To achieve this, the researcher calculates the cumulative probability of each residual using the formula:

$$P(i-th\ residual) = \frac{i}{(N+1)}$$

Where P is the cumulative probability of an observed residual, I is residual observation and N is the number of observations made within the given time period.

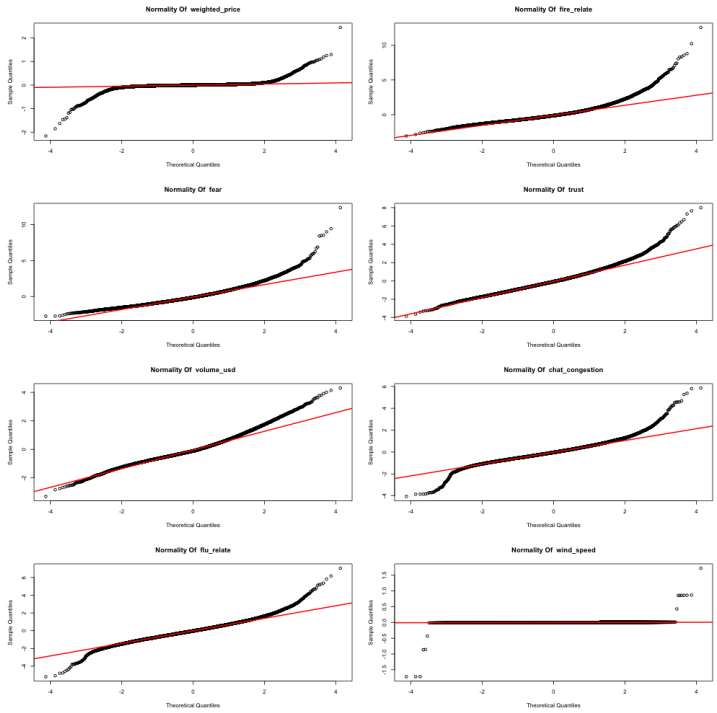

Figure 5-12: Normality Probability Plot of Residuals (Antecedents Phase)

The normal probability plot identifies departures from normality, non-linearity and outliers in a given distribution. Deviations from the straight line indicate a deviation from normality. A homoscedastic distribution is observed across all plots with reducing errors around zero. The S-Shape formed by data points in the plots indicate a shift towards normality. However, a few deviations from the 'normality' curve are observed in all 8 features. Given that the residuals were 'normal enough' with at least 98% of the data points falling within ±3 standard deviations of the mean, the effects of these outliers in the residuals were not significant.

The performance of the VECM model is measured using structural multivariate time series model validation techniques outlined in the literature review. The model is validation using both in sample and out of sample prediction values. The in-sample residuals are derived by obtaining the differences between the actual observations and the predictions returned by the model. The results for the in-sample model performance is shown in the table below.

	MAE	RMSE	MPE	MAPE	MAD	FEATURE ACCURACY
Weighted Price	0.03	0.08	-1.66	1.27	0.00	97.76
Fire Relatedness	0.65	0.90	-114.63	12.74	0.00	69.06
Fear	0.73	0.95	-4.34	334.60	0.00	66.46
Trust	0.74	0.96	-172.39	188.90	0.00	65.99
Volume	0.56	0.73	11.57	89.50	0.00	69.52
Chatroom Congestion	0.46	0.61	-975.04	221.84	0.00	78.71
Flu Relatedness	0.59	0.78	-1802.55	546.85	0.00	72.51
Average Wind Speed	0.00	0.03	-0.06	0.05	0.00	99.44
AVERAGE TOTAL	**0.47**	**0.63**	**-382.39**	**9.47**	**0.00**	

Table 5-12: In Sample Model Performance Results (Antecedents Phase)

The model works well in predicting most of the features in the model except the Chat Congestion and Volume features which records 22% and 3067% errors in total respectively. The model also performed accurately for features with smaller values of RMSPE, MAD and MSPE. These results show that the model works well with expected data and fits the observations of interest.
Additionally, out of sample validations are also done for each feature Out of sample forecasts are obtained by forecasting the N-ahead data values where N=60 minutes i.e the last hour initially subtracted from the original dataset. The out of sample residuals are then derived by obtaining the differences between forecasts on the test datasets using the VECM model and the test data with 60 observations. The measures of out of sample performance values are presented in the table below.

	MAE	RMSE	MPE	MAPE	MAD	FEATURE ACCURACY
Fire Relatedness	0.73	1.00	-73.25	26.92	0.00	65.39
Fear	0.78	0.99	-96.74	6.39	0.00	64.48
Trust	0.79	1.01	-108.36	19.67	0.00	64.04
Volume	0.79	1.01	-97.29	6.62	0.00	63.91
Chatroom Congestion	0.68	1.03	-99.09	4.67	0.00	65.83
Flu Relatedness	0.79	1.00	-93.06	3.80	0.00	64.10
Average Wind Speed	1.04	1.05	-104.37	11.21	0.00	58.28
AVERAGE TOTAL	0.84	1.06	-85.12	1.55	0.00	55.0

Table 5-13: Out of Sample Model Performance Results (Antecedents Phase)

The out of sample performance shows that the model works well in predicting new values of the endogenous feature. The Chat Congestion and Volume features are still seen to be uncaptured by the model. Although these features are selected as being important for predicting the outcome variable, the features included in the model are not entirely efficient in predicting them. By averaging the performance measure across the various performance measurement techniques, the model records a total of 85% in sample accuracy and 70% out of sample accuracy for the outcome feature.
Following the performance test, the granger causality tests each feature combination with the optimum lag selected in the model order selection stage. The structural model can also be used to make inference about the direction or directions of causality between every pair of features in the feature set. The granger causality test is set up with the null hypothesis of no causality between the feature ‘x’ (on the left-hand of the arrow) and the feature ‘y’ (on the right-hand of the arrow).
The table below shows only causal links with significant P-values to reject the null hypothesis.

SN	Causal Relation	F-Statistic	P-value
1.	**Fire Relatedness --> Weighted Price**	10.77	0.00
2.	**Fire Relatedness --> Fear**	35.89	0.00
3.	**Fire Relatedness --> Trust**	73.81	0.00
4.	**Fire Relatedness --> Chatroom Congestion**	10.80	0.00
5.	**Fire Relatedness --> Flu Relatedness**	79.63	0.00
6.	**Fire Relatedness --> Average Wind Speed**	15.10	0.00
7.	**Fear --> Weighted Price**	40.26	0.00

8.	**Fear --> Fire Relatedness**	23.46	0.00
9.	**Fear --> Trust**	16.39	0.00
10.	**Fear --> Chatroom Congestion**	63.69	0.00
11.	**Fear --> Flu Relatedness**	6.75	0.01
12.	**Trust --> Weighted Price**	7.93	0.00
13.	**Trust --> Fire Relatedness**	21.59	0.00
14.	**Trust --> Fear**	13.90	0.00
15.	**Trust --> Flu Relatedness**	196.12	0.00
16.	**Trust --> Average Wind Speed**	14.00	0.00
17.	**Volume --> Weighted Price**	22.13	0.00
18.	**Volume --> Flu Relatedness**	10.91	0.00
19.	**Volume --> Average Wind Speed**	75.13	0.00
20.	**Chatroom Congestion --> Weighted Price**	7.80	0.01
21.	**Chatroom Congestion --> Fear**	20.56	0.00
22.	**Chatroom Congestion --> Trust**	5.90	0.02
23.	**Chatroom Congestion --> Flu Relatedness**	94.79	0.00
24.	**Flu Relatedness --> Weighted Price**	103.58	0.00
25.	**Flu Relatedness --> Fire Relatedness**	25.29	0.00
26.	**Flu Relatedness --> Trust**	105.28	0.00
27.	**Flu Relatedness --> Volume**	12.26	0.00
28.	**Flu Relatedness --> Chatroom Congestion**	147.05	0.00
29.	**Flu Relatedness --> Average Wind Speed**	12.20	0.00
30.	**Average Wind Speed --> Volume**	9.94	0.00

Table 5-14: F-Test Results for Granger Analysis (Antecedents Phase)

The table above shows pairs of significant causal relations. The granger analysis shows a total of 30 causal relations of 56 possible relations between the 8 features in the feature set. Emotion features, fear and trust, from microblogging feeds are seen to likely have a causal relation with at least 4 features in the feature set. There is an observed inter-dimensional casual relation between multiple sets of features on the social dimension as well as an intra-dimensional causal relation between features on the economic and social dimension.

5.4.2 Stage 2: The Reconnaissance Phase

Stage 2 of the experiment integrates evidences from the social and physical dimension to spot indicators of the first phase of the kill-chain: an active reconnaissance on a target network. The rationale for conducting this stage of the experiment is to establish entanglements between events on the social dimension and physical dimension of cyberspace by identifying early indicators of activities on the network layer using evidences from the social dimension. The researcher uses series from the network layer of the physical dimension and the social dimension in constructing a predictive model.

Several endogenous variables are selected at this stage to characterize the occurrence of an active reconnaissance using a target's network flow data.
Three features are derived at this stage to indicate an active recon: Recon (Bhuyan, Bhattacharyya and Kalita, 2011), Scanner and the connection fail ratio (Nam and Kim, 2006). The feature 'scanner', which targets the same port on several hosts or several ports on a single host, is derived as a function of the number of distinct destinations Ip addresses and distinct destination ports communicating on a network within a given time window ***T***.

$$\frac{Number\ of\ distinct\ Destination\ IP\ addresses - Number\ of\ distinct\ Source\ IP\ addresses}{Number\ of\ distinct\ destination\ ports}$$

Equation 5-13: Estimating Network Scanners

The numerator of this formula gives the number of unique hosts communicating with each node in a network at any given point in time. Scaling this value by the number of destination ports gives the number of unique services used by each incoming request to every node on a network.

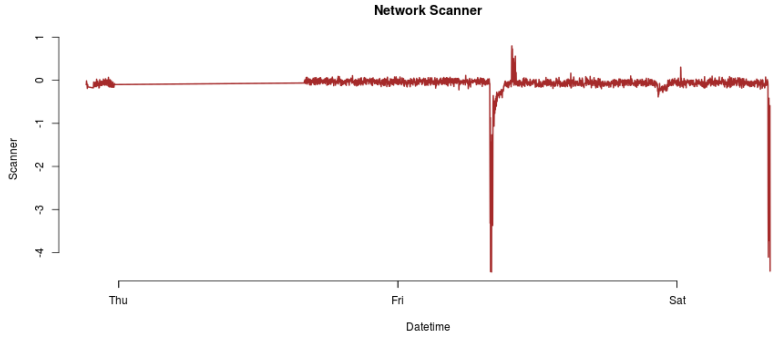

Figure 5-13: Network Scanners From Network Flow Data

Similarly, a value for 'Recon' within any given time period T is derived as a function of distinct node communicating on a network and the total number of connections to distinct nodes on a network. This characterization follows methods developed by Bailey, Roedel and Silenok's (Bailey Lee, Roedel and Silenok, 2003) as discussed in the literature review. For a given time window ***T***, the researcher estimates the value of 'Recon' as:

$$\frac{Number\ of\ distinct\ Source\ IP\ addresses\ attempting\ to\ connect\ to\ a\ source}{Number\ of\ conenction\ attempts\ to\ distinct\ IP\ addresses\ by\ a\ single\ source}$$

Equation 5-14: Estimating Network Reconnaissance

The numerator of this formula gives the distinct nodes within a network communicating with each distinct destination ip addresses. The denominator gives the total number of For normal traffic, for any given time frame, the value of this figure should be as close to the mean or 1 as possible. On the other hand, the researcher also estimates values for a 'vertical scan' within a given time window T.

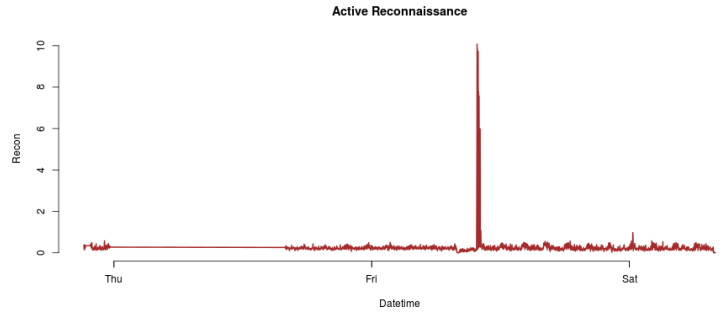

Figure 5-14: Active Reconnaissance from Network Flow Data

Finally, (Kinable, 2008) presents a hypothesis that measures the connection fail ratio of a network at any given point in time as a function of the amount of inbound connections and outbound connections. The CFR of a network is given as:

$$\frac{Number\ of\ outbound\ connections - Number\ of\ inbound\ connections}{Number\ of\ outbound\ connections}$$

Equation 5-15: Estimating Network Connection Fail Ratio

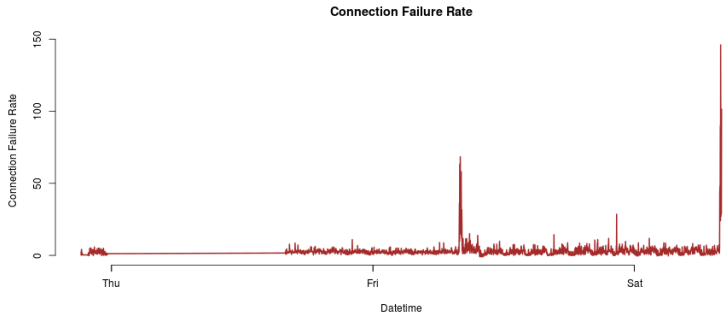

Figure 5-15: Connection Fail Ratio

The endogenous variable selected at this stage is the 'Recon' which indicates the occurrence of an active reconnaissance on a target network. Datasets from both dimensions were integrated on a similar timeline with the aim of predicting events of the last hour on the physical dimension. The merging of both datasets produced a single dataset with 69 features and 2524 observations. Additionally, the data was split into two removing the last 60 observations (the last hour) from each feature in the dataset. The out of sample reliability of the constructed model will be tested on observations from the last hour using an N-ahead prediction strategy where N=60. This is done to ensure reliability of out of sample prediction performance testing.

A simple intra-dimensional correlation analysis of the 69 features on both dimensions as shown in the figure below, reveals a slight correlation between features on both dimensions. For example, the average wind speed on the social dimension is seen to have a strongly positive correlation coefficient with several features on the physical dimension. The black boxes in the figure below groups features into hierarchical clusters. The recon, connection fail ratio and scanner features are seen to be highly correlated with other network layer features such as ip addresses, ports, number of connections denied

by intrusion detection system and total alerts raised by intrusion detection system. Emotion and opinion features are also seen to be highly correlated with each other on the social dimension.

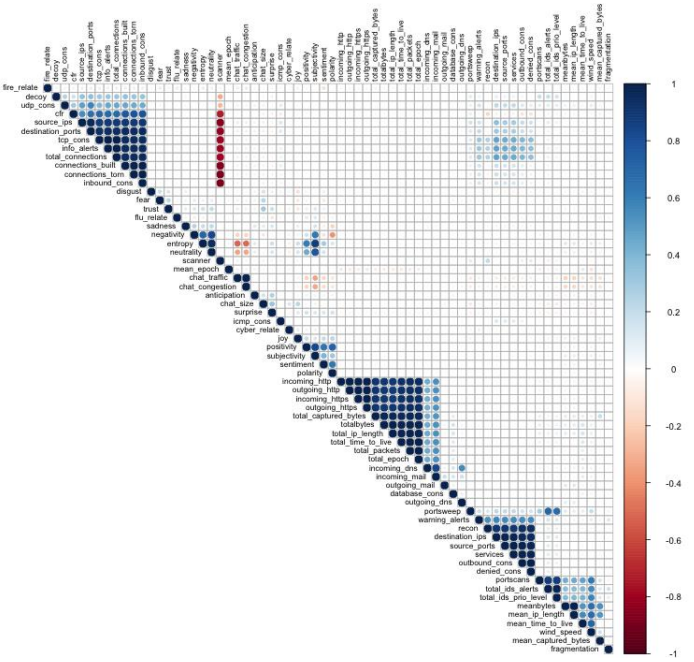

Figure 5-16: Intra-Dimensional Correlation Coefficients (Reconnaissance Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment. The correlation coefficients are on a scale of +1 to -1 as shown by the colour bar. Highly positively correlated pairs of variables tend towards +1 while highly negatively correlated pairs of variables tend towards -1. Multiple pairs of variables are seen to be highly positively or negatively correlated. For example, the group of features, Cyber relatedness ('cyber_relate'), Entropy, chat room congestion ('chat_congestion') and chat room traffic ('chat_traffic') posted per minute are all seen to be highly positively correlated. Additionally, the group of features such as packet data features (totalBytes, total_packets, total_captured_bytes, incoming_http, outgoing_http) and intrusion detection features (total_ids_alerts, total_ids_prio_level, fragmentation) are also seen to be highly correlated. Consequently, one or more of pairs of features that are seen to be highly correlated with each other (where Pearson's correlation coefficient $\geq +0.65$ or ≤ -0.65) are removed from the dataset. The correlation elimination process reduces the feature set by removing one or more of a set of highly correlated variables (Hall and Smith, 1998). The assumption is that information provided by highly correlated variables can be provided to the model by a single one of those variables therefore reducing redundancy to achieve a parsimonious model. The results of the filtered feature set after a correlation elimination produces a smaller feature set of 30 features as shown in the figure below, eliminating 36 highly correlated features.

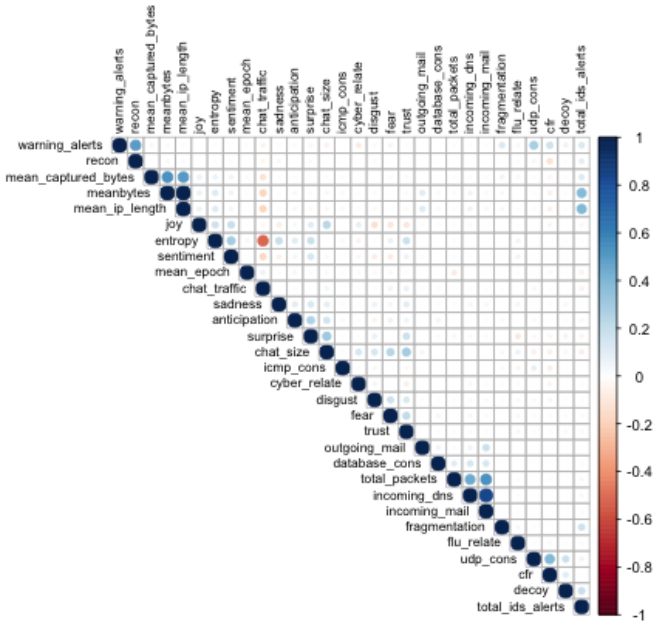

Figure 5-17: Correlation Co-efficient of Reduced Feature set (Reconnaissance Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment for the reduced feature set. This reduced set of 30 variables including the endogenous variable, 'recon' is selected for further feature selection methods. To further reduce the number of features, the researcher applies a recursive features selection method to select only features that significantly reduce the error of prediction in the endogenous feature. The recursive feature selection stage further reduces the uncorrelated feature set to only those variables that minimizes the prediction error for the chosen endogenous variable 'weighted_price'. To begin the recursive feature selection from the uncorrelated feature set, the researcher starts by building an ordinary least squares regression linear model with the 'recon' as the endogenous feature predicted by all other features in the reduced feature set above as shown in the figure above. The recursive feature selection searches for a model that optimizes the Akaike Information (AIC) and reduces errors. The recursive feature selection step further reduces the number of features leaving 10 features. The resulting features are tabled below and used in further analysis. Lastly, the researcher creates generic methods for calculating the importance of each feature in predicting the outcome feature. This is used to examine the contribution of each feature in the dataset in predicting the outcome feature 'recon'. This last step in the feature selection phase is elimination by variable importance (Mehmood *et al.*, 2012). The absolute value of the t-statistic for each feature is used to compute the unique contribution of each input-feature to the model and a cut off of 1 is chosen (Noppamas, Seree and Kidakan, 2014). From the figure below, it is observed that the most important features for predicting the outcome feature 'recon' are the connection fail ratio, cyber-relatedness ('cyber_relate') and flu-relatedness of microblogging feeds ('flu_relate'), emotion features from microblogging feeds such as sadness and surprise, total number of alerts ('total_ids_alerts') raised by intrusion detection system, total number of UDP connections ('UDP') and total number of warning alerts ('warning_alerts') by intrusion detection system.

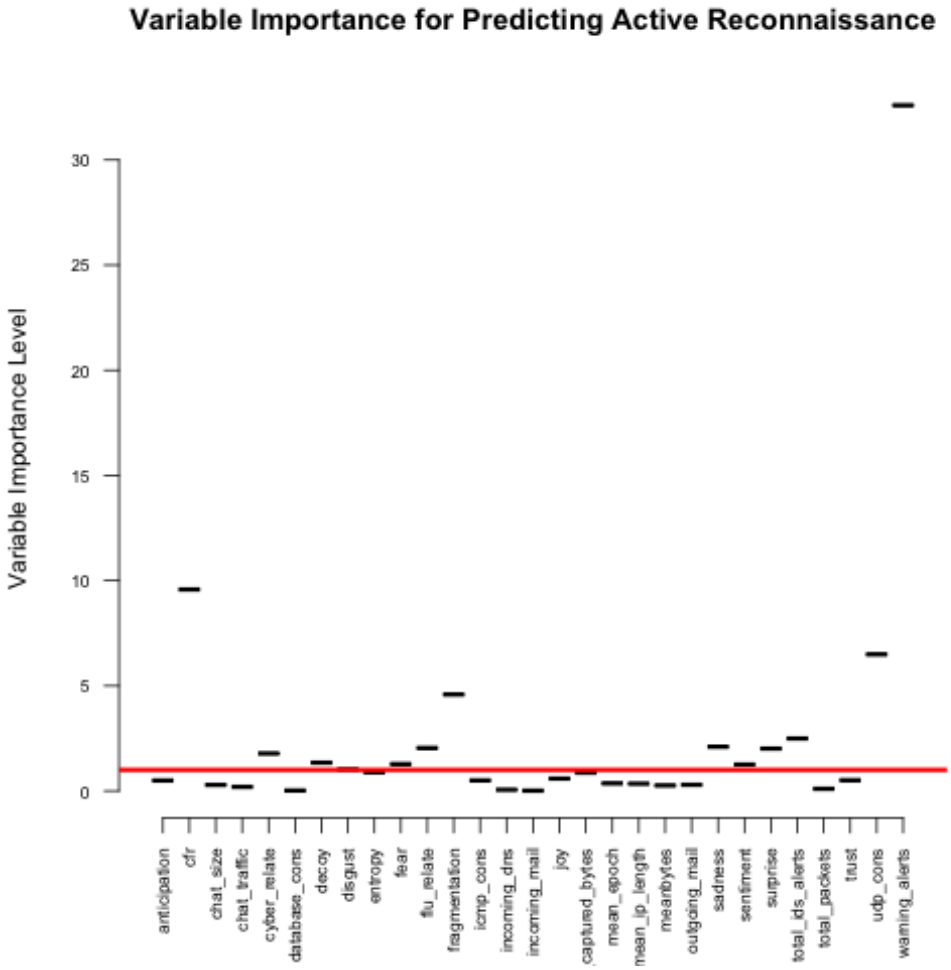

Figure 5-18: Variable Importance Of Features (Reconnaissance Phase)

The last phase of the feature selection left 11 homogenous features as extremely useful for predicting the outcome feature. The final features for the model are shown in table 5-14 below. Intrusion detection features such as the connection fail ratio ('cfr'), number of fragmentation alerts ('fragmentation') and the number of warnings raised ('warning') are seen to be important during the reconnaissance phase of the kill-chain. These features are indicative of adversary presence in a target's network. Additionally, on the social dimension, emotion-based features (fear, sadness and surprise) are also indicators to monitor at the reconnaissance phase.

The next phase of the analytical framework involves selecting appropriate orders for the intended model. This involves selecting an autoregressive order, testing for seasonality in each feature and selecting a moving average as discussed in the analytical framework.

The VAR model order selection phase seeks to select the appropriate significant lag at which features are serially correlated. The researcher uses an iterative solution with the maximum of 100 lags. At each iteration, the researcher builds a model for each equation in the VAR and measures the AIC. Finally, the researcher selects 24 lags (approximately half-hour) as this is the lag length with the optimal AIC. The partial auto correlation plots below show the significant lags each of the ten features selected.

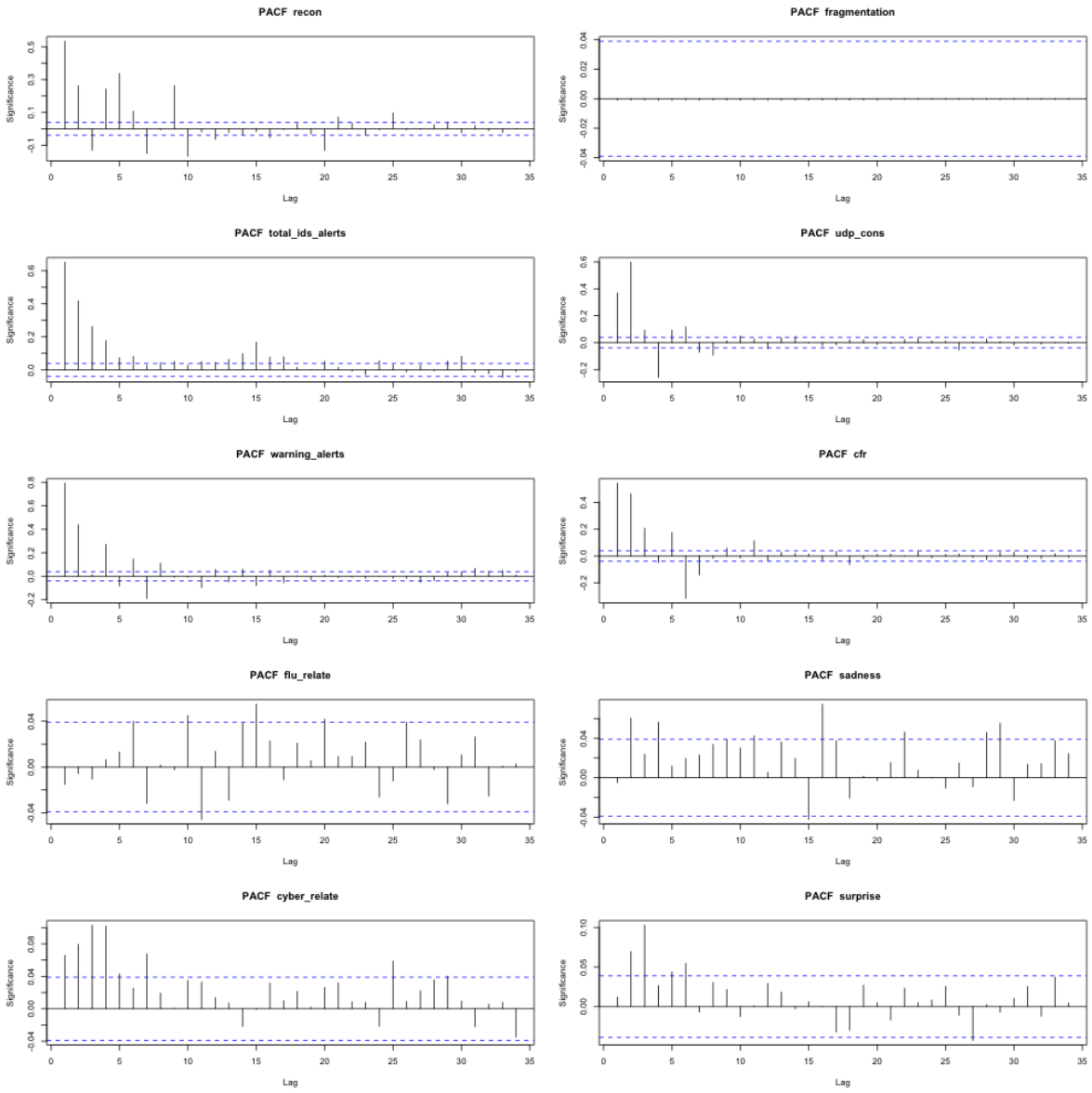

Figure 5-19: Partial Auto Correlation of Selected Features (Reconnaissance Phase)

In addition to selecting appropriate VAR lag order, the researcher also tests each time series in the feature set for seasonality. In order to achieve this, the researcher employs a time series decomposition technique. As discussed in the literature review, time series decomposition works by splitting a time series into its three main components: the trend, the seasonal movement and the error terms. Seasonal patterns are expected to repeat with a fixed period of time. An additive decomposition method was chosen as the magnitude of the series does not increase with the series. After applying the Fourier transform for detecting seasonality, seven of the ten features exhibited seasonal trend. Features on the network layer such as total number of IDS alerts, recon, and connection fail ratio and number of UDP connections were seen to have between 5- and 10-hours seasonal variations. Emotion probabilities such as 'Sadness' and 'Surprise' have a 12-hour seasonal variation. Additionally, the cyber relatedness and flu-relatedness of texts in microblogging feeds are seen to have a 12-hour seasonal time period.

A stationarity and normality test are conducted on the 10 features and results are tabled below. The results of the stationarity test show a significant p-value for rejecting the null hypothesis of non-

stationarity in the augmented dickey fuller stationarity test. All selected features are seen to have a constant mean and variance over the selected period of time.

	Stationarity Test			Normality Test	
	P-value	**Augmented Dickey Fuller Test (Test Statistic)**	**Stationary**	**Skewness**	**Kurtosis**
Recon	0.01	-9.23408	True	16.0768	282.1766
Fragmentation	0.01	-13.4934	True	35.4824	1257.001
Total Ids Alerts	0.01	-5.60466	True	2.647027	12.87255
Udp Connections	0.01	-6.30005	True	8.676738	131.6894
Warning Alerts	0.01	-6.5179	True	7.673929	62.5321
Connection Fail Ratio	0.01	-5.33053	True	13.7233	251.8265
Flu Relatedness	0.01	-13.1509	True	0.235649	1.225698
Sadness	0.01	-10.9857	True	1.447701	8.616825
Cyber Relatedness	0.01	-10.6626	True	0.683449	1.654992
Surprise	0.01	-11.4438	True	1.286536	6.798069

Table 5-15: Features' Stationarity and Normality Test Results (Reconnaissance Phase)

Given that all features were stationary, the Johansen's test for co-integration was not applicable at this stage of the experiment. However, for further investigation, the Engle and Granger test for co-integration was used to determine co-integration between each pair of features in the feature set. To achieve this, the researcher constructs a linear regression model between each pair of features and conducts a stationarity test on the residuals of the resulting model. The Engle Granger test shows significant p-values for three features: the total number of IDS Alerts on the physical dimension, the level of population sadness and the level of cyber relatedness of microblogging feeds on the social dimension derived from microblog feeds. The figure below shows a long-run equilibrium relationship between these three features with this relationship strongest towards the end of the period of observation.

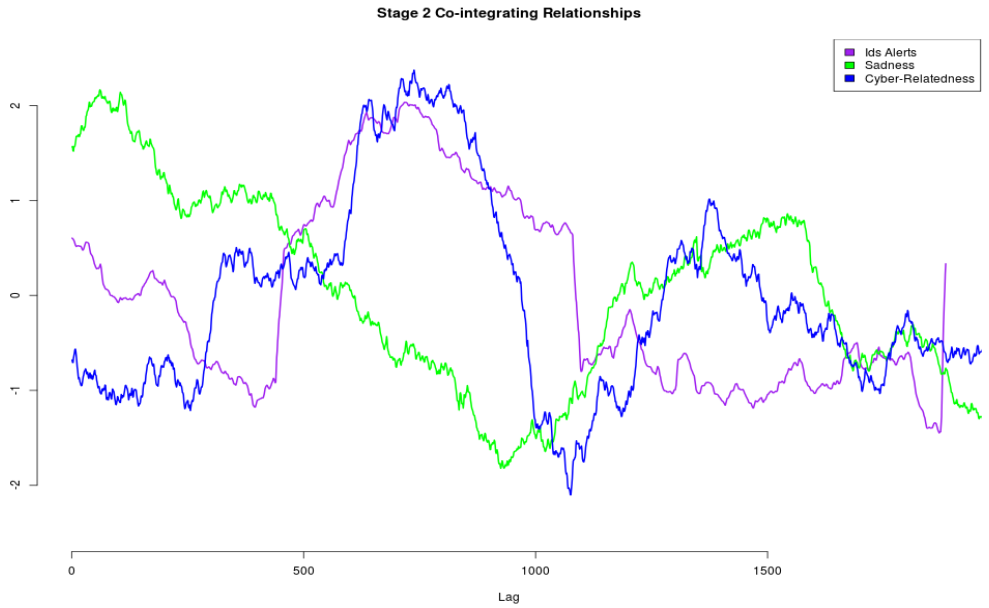

Figure 5-20: Co-Integrating Relationships Identified in the Reconnaissance Phase

The figure above visually represents the identified co-integrating relationships between the total number of IDS alerts, sadness as determined by emotion detectors and the cyber-relatedness of users' micro-blogging texts. This identified link between the social and physical dimensions at the reconnaissance phase of the kill-chain. Although, these co-integrating relationships do not indicate a 'causal' link at this point, the possibilities of this connections highlight the need to monitor these features at the reconnaissance phase. The methods presented here aere able to pick up these entanglements between events on the social dimension that led to events on the physical dimension as designed in the scenario.

Given that all features are observed to be stationary as seen in table 5-15 below, a VAR model was fitted to the levels. The VAR with order of 24 lags was estimated utilizing the OLS per equation in the model with 10 parameter OLS models where each feature is predicted by its on lags and lags of the other 9 features in the model. The density distribution plots below shows the residuals from the fitted VAR(24) model. The error terms i.e. the residuals are expected to be independently distributed across each equation and serially uncorrelated. The residuals are experimental errors derived by finding the difference between the observed data points and the predicted data points. To ensure that these assumptions are met, the researcher tests the hypothesis of white noise residuals using a normality test on the residuals of the VAR model. The stationarity and normality test results for the residuals are shown below.

	Stationarity Test			**Normality Test**	
	P-Value	**Augmented Dickey Fuller Test (Test Statistic)**	**Stationary**	**Skewness**	**Kurtosis**
Recon	0.01	-12.10	True	8.17	248.24
Fragmentation	0.01	-13.52	True	7.32	155.97
Total Ids Alerts	0.01	-13.98	True	2.02	32.12
Udp Connections	0.01	-15.22	True	0.22	1.19

Warning Alerts	0.01	-14.34	True	2.59	38.56
Connection Fail Ratio	0.01	-13.66	True	5.88	140.65
Flu Relatedness	0.01	-14.68	True	0.52	1.05
Sadness	0.01	-14.84	True	1.01	4.43
Cyber Relatedness	0.01	-12.10	True	8.17	248.24
Surprise	0.01	-13.52	True	7.32	155.97

Table 5-16: Residual Stationarity and Normality Test Results (Reconnaissance Phase)

The distribution of the residuals for intrusion detection features are seen to be centered on zero with narrow peaked curves. The model performs moderately in capturing the trends in the other features as the residuals are seen to be slightly skewed to the left.

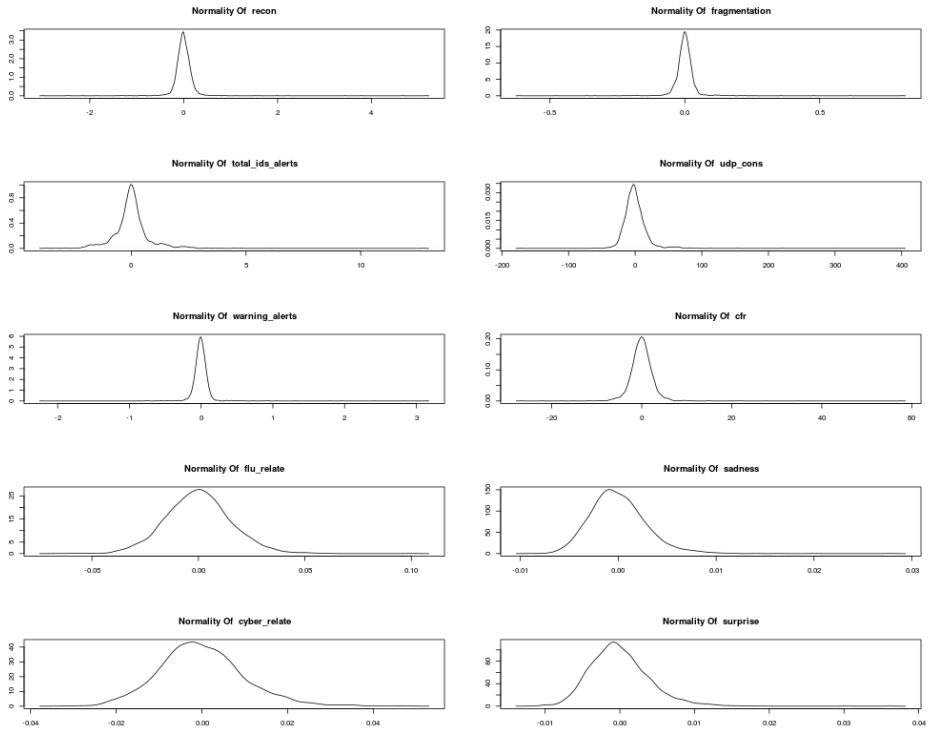

Figure 5-21: Density Distribution of Residuals (Reconnaissance Phase)

Lastly, after observing the density distribution of the residuals, the researcher derives a normal probability plot for each residual from the structural model. To achieve this, the researcher calculates the cumulative probability of each residual using the formula:

$$P(i-th\ residual) = \frac{i}{(N+1)}$$

Equation 5-16: Residual Cumulative Probability

Where P is the cumulative probability of an observed residual, I is residual observation and N is the number of observations made within the given time period.

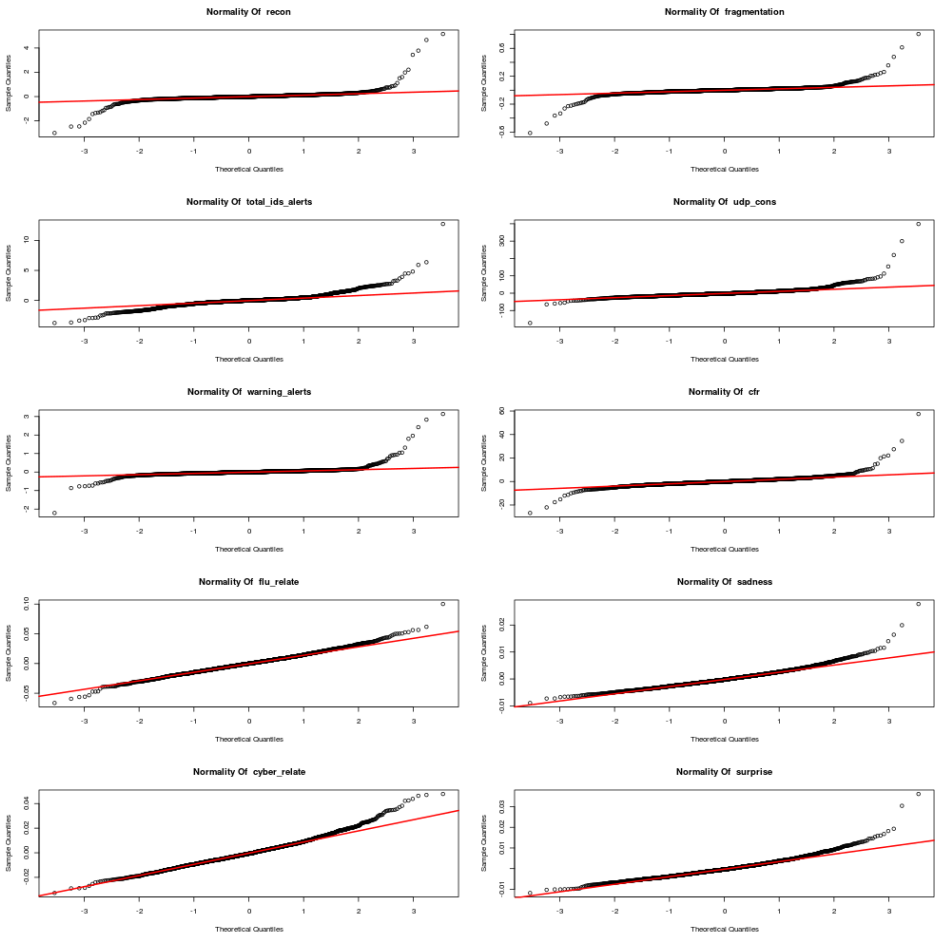

Figure 5-22: Normality Probability Plot of Residuals (Reconnaissance Phase)

The normal probability plot identifies departures from normality, non-linearity and outliers in a given distribution. Deviations from the straight line indicate a deviation from normality. A homoscedastic distribution is observed across all plots with reducing errors around zero. The S-Shape formed by data points in the plots indicate a shift towards normality. However, a few outliers are observed in all 10 features.

The performance of the VAR model is measured using structural multivariate time series model validation techniques outlined in the analytical framework. The model is validation using both in sample and out of sample prediction values. The in-sample residuals are derived by obtaining the differences between the actual observations and the predictions returned by the model. The results for the in-sample model performance is shown in the table below.

	MAE	RMSE	MPE	MAPE	MAD	FEATURE ACCURACY
Recon	0.26	0.57	-86.20	86.20	0.00	79.35%
Fragmentation	0.13	0.26	-5.05	5.05	0.00	90.28%
Total Ids Alerts	0.35	0.56	-17.31	17.31	0.00	77.29%
Udp Connections	0.35	0.59	-71.62	71.62	0.00	76.44%
Warning Alerts	0.17	0.39	53.13	53.13	0.00	59.57%
Connection Fail	0.31	0.52	81.34	81.34	0.00	38.54%

Ratio						
Flu Relatedness	0.73	0.94	840.32	840.32	0.00	-362.00%
Sadness	0.71	0.93	5069.51	5069.51	0.00	-2475.63%
Cyber Relatedness	0.71	0.92	123.91	123.91	0.00	-2.82%
Surprise	0.71	0.94	-3022.03	3022.03	0.00	58.71%
Average Model Error	0.44	0.66	296.60	937.04	0.00	

Table 5-17: In Sample Model Performance Results (Reconnaissance Phase)

The model works well in predicting most of the features in the model except the UDP connections feature which records a 637% average error. The model also performed accurately for features with smaller values of RMSPE, MAD and MSPE. These results show that the model works well with expected data and fits the observations of interest.

Additionally, out of sample validations are also done for each feature Out of sample forecasts are obtained by forecasting the N-ahead data values where N=60 minutes i.e the last hour initially subtracted from the original dataset. The out of sample residuals are then derived by obtaining the differences between forecasts on the test datasets using the VAR model and the test data with 60 observations. The measures of out of sample performance values are presented in the table below.

	MAE	**RMSE**	**MPE**	**MAPE**	**MAD**	**FEATURE ACCURACY**
Recon	0.21	0.45	-281.27	281.27	0.01	62.14
Fragmentation	0.00	0.00	0.00	0.00	0.00	0.00
Total Ids Alerts	0.00	0.00	0.00	0.00	0.00	0.00
Udp Connections	0.20	0.11	-217.71	217.71	0.01	65.71
Warning Alerts	0.33	0.45	-109.40	109.40	0.01	77.79
Connection Fail Ratio	0.24	0.33	-197.82	197.82	0.01	78.80
Flu Relatedness	0.55	1.54	-75.23	75.23	0.01	53.17
Sadness	0.22	1.47	-193.78	193.78	0.01	52.45
Cyber Relatedness	0.32	0.32	-110.55	110.55	0.01	55.35
Surprise	0.44	0.11	-74.10	74.10	0.01	53.54
Average Model Error	0.25	0.48	296.60	937.04	0.00	

Table 5-18: Out of Sample Model Performance Results (Reconnaissance Phase)

The out of sample performance shows that the model works well in predicting new values of the endogenous feature. However, some features such as predictions for the UDP connections and the level of cyber relatedness of microblogging feeds are seen to exhibit a higher level of out of sample percentage error. Although these features are selected as being important for predicting the outcome variable, the features included in the model are not entirely efficient in predicting them. By averaging the performance measure across the various performance measurement techniques, the model records a total of 85% in sample accuracy and 67% out of sample accuracy for the outcome feature.

Following the performance test, the granger causality tests each feature combination with the optimum lag selected in the model order selection stage. The structural model can also be used to make inference about the direction or directions of causality between every pair of feature in the

feature set. The granger causality test is set up with the null hypothesis of no causality between the feature 'x' (on the left-hand of the arrow) and the feature 'y' (on the right-hand of the arrow).
The table below shows only causal links with significant P-values to reject the null hypothesis.

SN	Causal Relationship	F Statistic	P-value
1.	**Reconnaissance --> Ids Alerts**	18.54	0.00
2.	**Reconnaissance --> Warning Alerts**	233.30	0.00
3.	**Fragmentation --> Warning Alerts**	23.37	0.00
4.	**Ids Alerts --> Reconnaissance**	7.61	0.01
5.	**IDS Alerts --> Fragmentation**	32.03	0.00
6.	**Ids Alerts --> Warning Alerts**	5.47	0.02
7.	**Udp Connections --> Ids Alerts**	6.95	0.01
8.	**Udp Connections --> Warning Alerts**	25.97	0.00
9.	**Udp Connections --> Connection Fail Ratio**	688.54	0.00
10.	**Warning Alerts --> Reconnaissance**	121.54	0.00
11.	**Warning Alerts --> Ids Alerts**	10.79	0.00
12.	**Warning Alerts --> Udp Connections**	5.96	0.01
13.	**Warning Alerts --> Connection Fail Ratio**	36.85	0.00
14.	**Connection Fail Ratio --> Udp Connections**	1033.82	0.00
15.	**Connection Fail Ratio --> Warning Alerts**	67.41	0.00
16.	**Sadness --> Cyber Relatedness**	5.33	0.02
17.	**Sadness --> Surprise**	14.87	0.00
18.	**Cyber Relatedness --> Udp Connections**	5.81	0.02
19.	**Cyber Relatedness --> Warning Alerts**	12.02	0.00
20.	**Cyber Relatedness --> Sadness**	7.45	0.01

Table 5-19: F-Test Results for Granger Analysis (Reconnaissance Phase)

The granger analysis shows a total of 20 likely causal relations of 100 possible relations between the 10 features in the feature set. The recon feature is seen to have plausible causal relations with 2 intrusion detection features. There is also an observed bi-causal relation or feedback loop between the recon feature and these two features. Similarly, the connection fail ratio and the number of UDP connections are also seen to have a bi-directional causal relationship. Inter-dimensionally, there exist causal relationships between features on the physical dimension as well as causal relations between features across the social and physical dimension, intra-dimensionally.

5.4.3 Stage 3: The Weaponization Phase

Stage 3 of the experiment integrates evidences leading from the physical dimension and evidences on the social dimension to spot indicators of a plausible ongoing weaponization on the social dimension of cyberspace. The rationale for conducting this stage of the experiment is to establish entanglements between events on the physical and social dimension of cyberspace using co-integral links between them. The researcher uses series from the network layer of the physical dimension taking into account prior stages of the kill-chain and the social dimension in constructing a predictive model.

The endogenous variable selected for this phase characterizes the spread of discussions related to cyber weapons and vulnerabilities in microblogging feeds. The 'weaponizaton' variable is derived by quantifying each feed in microblogging feeds to an external weaponization 'Lexicon' or wordlist. The quantity is a bag-of-words n-gram representation of the number of words in the wordlist found in the feed.

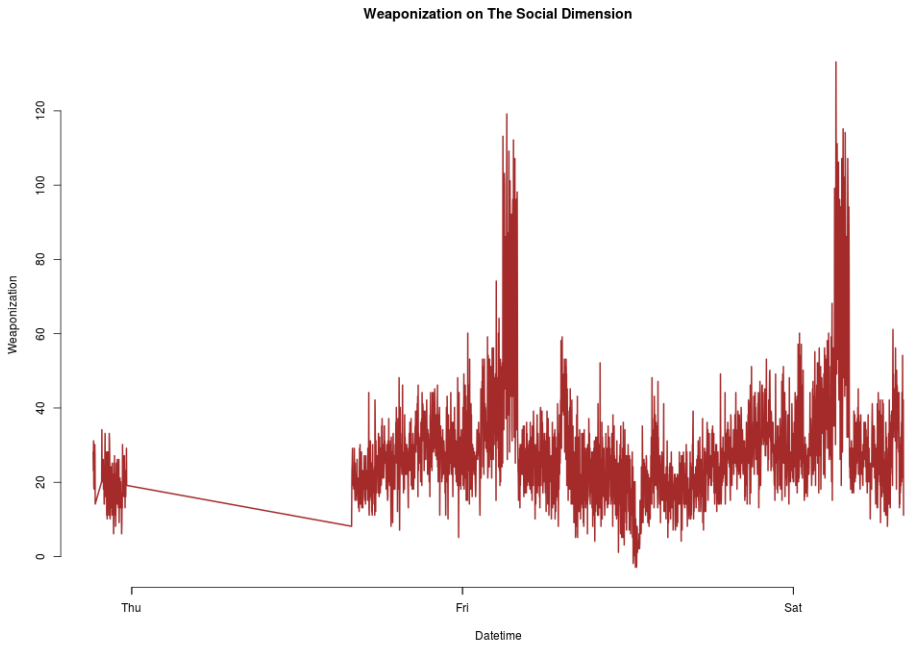

Figure 5-23: Characterizing the Weaponization Phase

The endogenous variable selected at this stage is the 'weaponization' which indicates discussion about cyber weapons, exploits and vulnerabilities on social platforms. Datasets from both dimensions were integrated on a similar timeline with the aim of predicting events of the last hour on the physical dimension. The merging of both datasets produced a single dataset with 71 features and 2404 observations. Additionally, the data was split into two removing the last 60 observations (the last hour) from each feature in the dataset. The out of sample reliability of the constructed model will be tested on observations from the last hour using an N-ahead prediction strategy where N=60. This is done to ensure reliability of out of sample prediction performance testing.

A simple intra-dimensional correlation analysis of the 71 features on both dimensions as shown in the figure below, reveals a slight correlation between features on both dimensions. For example, the level of chat congestion and chatroom traffic is seen to have a slight correlation with the number of distinct connections to the network. Additionally, on the social dimension, emotion and opinion features are observed to be highly correlated with each other. The black boxes in the figure below groups features into hierarchical clusters. The recon, connection fail ratio and scanner features are seen to be highly

correlated with other network layer features such as ip addresses, ports, number of connections denied by intrusion detection system and total alerts raised by intrusion detection system. This correlation was also observed in the previous stage of the experiment.

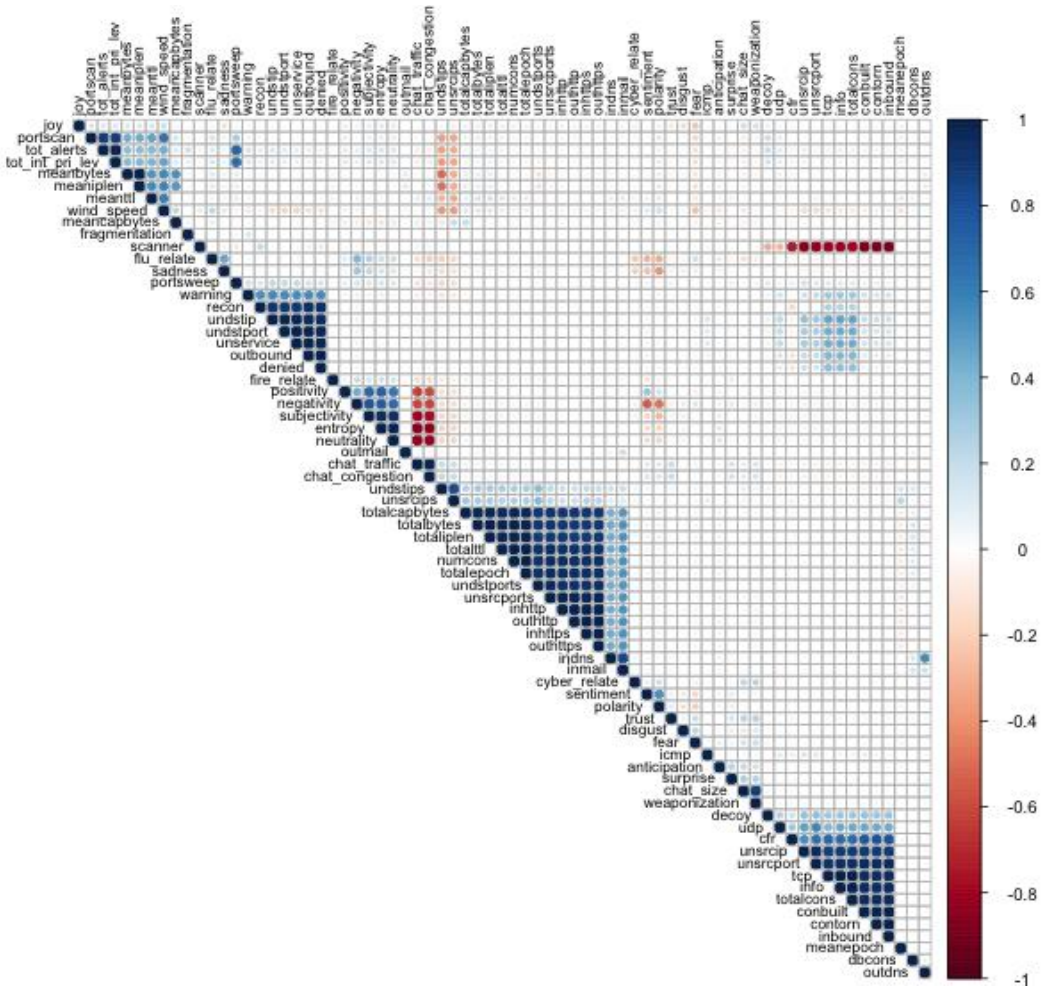

Figure 5-24: Intra-Dimensional Correlation Coefficients (Weaponization Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment. The correlation coefficients are on a scale of +1 to -1 as shown by the color bar. Highly positively correlated pairs of variables tend towards +1 while highly negatively correlated pairs of variables tend towards -1. Multiple pairs of variables are seen to be highly positively or negatively correlated. For example, the group of features, Cyber relatedness ('cyber_relate'), Entropy, chat room congestion ('chat_congestion') and chat room traffic ('chat_traffic') posted per minute are all seen to be highly positively correlated. Additionally, the group of features such as packet data features (totalBytes, total_packets, total_captured_bytes, incoming_http, outgoing_http), reconnaissance features ('cfr', 'recon', 'portscan','pingsweep') and intrusion detection features (total_ids_alerts, total_ids_prio_level, fragmentation) are also seen to be highly correlated with each other. There is also an observed subtle correlation between the 'weaponization' and event-based cyber-relatedness of user discussions 'cyber_relate'. Consequently, one or more of pairs of features that are seen to be highly correlated with each other (where Pearson's correlation coefficient $\geq +0.65$ or ≤ -0.65) are removed from the dataset. The correlation elimination process reduces the feature set by removing one or more of a set of highly correlated variables (Hall and Smith, 1998). The assumption is that information provided by highly correlated variables can be provided to the model by a single one of those variables therefore reducing redundancy to achieve a parsimonious model. The results of the

filtered feature set after a correlation elimination produces a smaller feature set of 25 features as shown in the figure below, eliminating 46 highly correlated features.

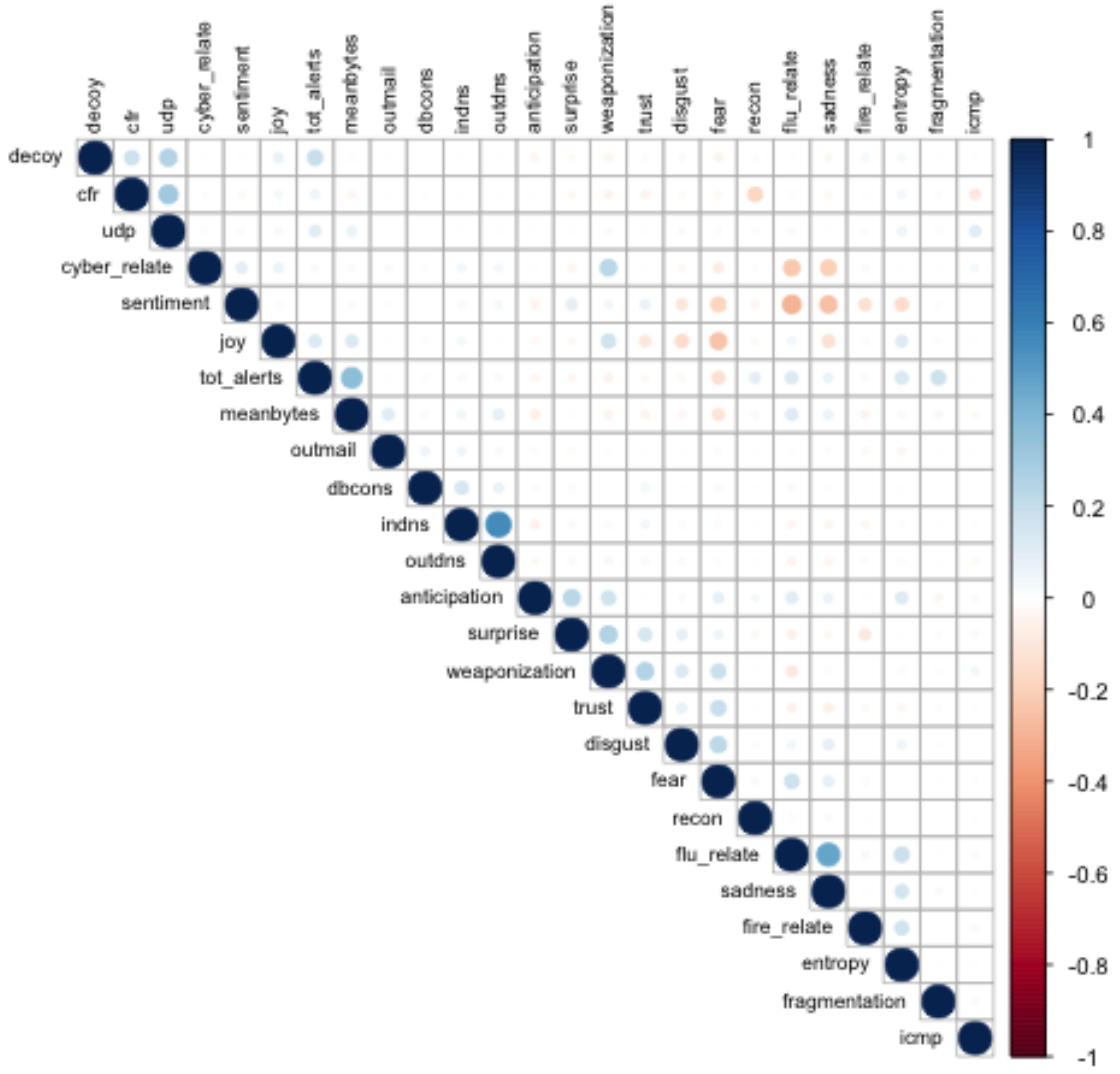

Figure 5-25: Correlation Co-efficient of Reduced Feature set (Weaponization Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment for the reduced feature set. This reduced set of 25 features including the endogenous variable, 'weaponization' are selected for further feature selection methods. To further reduce the number of features, the researcher applies a recursive features selection method to select only features that significantly reduce the error of prediction in the endogenous feature. The recursive feature selection stage further reduces the uncorrelated feature set to only those variables that minimizes the prediction error for the chosen endogenous variable 'weaponization'. To begin the recursive feature selection from the uncorrelated feature set, the researcher starts by building an ordinary least squares regression linear model with the 'recon' as the endogenous feature predicted by all other features in the reduced feature set above as shown in the figure above. The recursive feature selection searches for a model that optimizes the Akaike Information (AIC) and reduces errors. The recursive feature selection step further reduces the number of features leaving 15 features. The resulting features are shown in the table below and used in further analysis. Lastly, the researcher creates generic methods for calculating the importance of

each feature in predicting the outcome feature. This is used to examine the contribution of each feature in the dataset in predicting the outcome feature 'recon'. This last step in the feature selection phase is elimination by variable importance (Mehmood *et al.*, 2012). The absolute value of the t-statistic for each feature is used to compute the unique contribution of each input-feature to the model and a cut off of 1 is chosen (Noppamas, Seree and Kidakan, 2014). From the figure below, it is observed that the most important features for predicting the outcome feature 'Weaponization' are the flu relatedness, fire relatedness and cyber relatedness of microblogging feeds, emotion features from microblogging feeds such as population joy, fear, trust, total number of packet bytes transmitted.

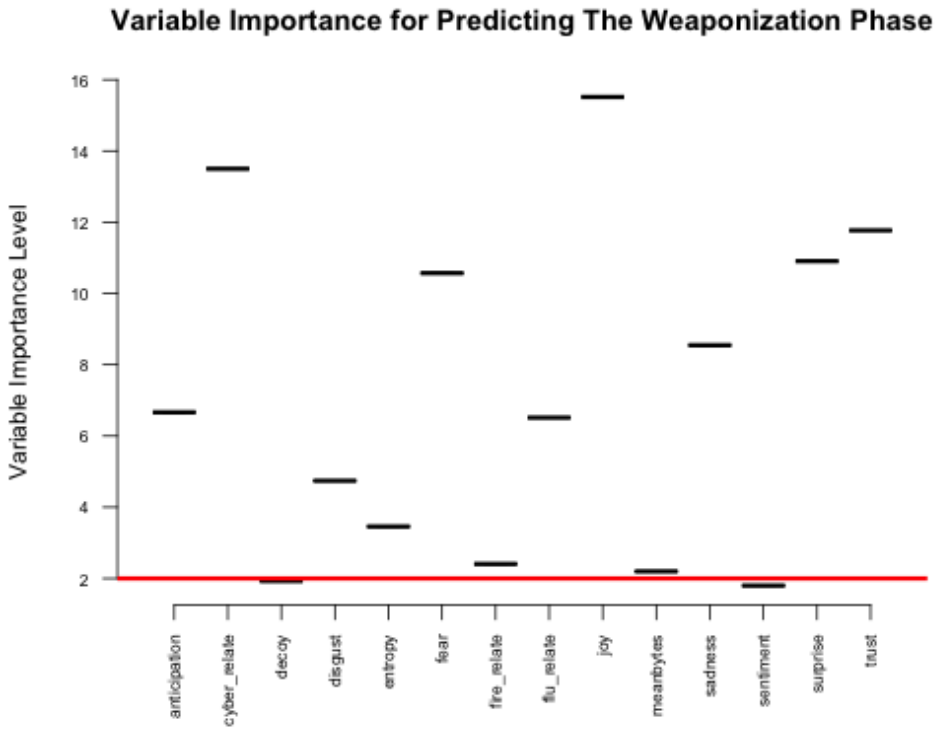

Figure 5-26: Variable Importance Of Features (Weaponization Phase)

The figure above shows the important indicators of the 'weaponization' phase of the kill-chain. Emotion-based features such as anticipation, disgust, sadness and trust are seen to be important indicators at this phase. Additionally, event-based features such as cyber, flu and fire relatedness (in the context of the scenario) are important indicators at this phase as well. The last phase of the feature selection left the same 15 homogenous features as extremely useful for predicting the outcome feature. The final features for the model are shown in the table 5-19 below.

The next phase of the analytical framework involves selecting appropriate orders for the intended model. This involves selecting an autoregressive order, testing for seasonality in each feature and selecting a moving average as discussed in the analytical framework.

The VAR model order selection phase seeks to select the appropriate significant lag at which features are serially correlated. The researcher uses an iterative solution with the maximum of 100 lags. At each iteration, the researcher builds a model for each equation in the VAR and measures the AIC. Finally, the researcher selects 5 lags (approximately 5 mins) as this is the lag length with the optimal AIC.

In addition to selecting appropriate VAR lag order, the researcher also tests each time series in the feature set for seasonality. In order to achieve this, the researcher employs a time series

decomposition technique. As discussed in the literature review, time series decomposition works by splitting a time series into its three main components: the trend, the seasonal movement and the error terms. Seasonal patterns are expected to repeat with a fixed period of time. An additive decomposition method was chosen as the magnitude of the series does not increase with the series. After applying the Fourier transform for detecting seasonality, none of the fifteen features exhibited seasonal trend at the selected time interval.

A stationarity and normality test is conducted on the 15 features and results are tabled below. The results of the stationarity test show a significant p-value for rejecting the null hypothesis of non-stationarity in the augmented dickey fuller stationarity test. All selected features are seen to have a constant mean and variance over the selected period of time.

	Stationarity Test			**Normality Test**	
	P-value	**Test Statistic**	**Stationary**	**Skewness**	**Kurtosis**
Weaponization	0.01	-7.68	True	0.40	0.96
Joy	0.01	-10.04	True	0.25	0.33
Mean Bytes	0.01	-7.29	True	1.92	8.49
Flu Relatedness	0.01	-7.07	True	0.47	0.61
Sadness	0.01	-8.89	True	0.49	0.50
Fire Relatedness	0.01	-7.05	True	1.85	6.73
Entropy	0.01	-6.75	True	1.14	3.46
Cyber Relatedness	0.01	-8.07	True	0.77	1.62
Sentiment	0.01	-9.27	True	0.55	-1.06
Trust	0.01	-8.60	True	0.43	1.14
Disgust	0.01	-12.27	True	1.72	5.05
Fear	0.01	-8.59	True	0.50	0.31
Anticipation	0.01	-10.00	True	0.65	0.57
Surprise	0.01	-9.02	True	0.55	1.60
Decoy	0.01	-8.54	True	15.41	235.40

Table 5-20: Features' Stationarity and Normality Test Results (Weaponization Phase)

Additionally, the skewness tests show that all features are at least fairly normal with varying peakedness shown by the kurtosis test. Given that all features were stationary, the Johansen's test for co-integration was not applicable at this stage of the experiment. However, for further investigation, the Engle and Granger test for co-integration was used to determine co-integration between each pair of feature in the feature set. To achieve this, the researcher constructs a linear regression model between each pair of feature and conducts a stationarity test on the residuals of the resulting model. The Engle Granger test shows no significant p-values for features in the system at the selected time period.

Given that all features are observed to be stationary and the absence of any co-integrating relationships at this stage of the experiment, a VAR model was fitted to the levels. The VAR with order of 5 lags was estimated utilizing the OLS per equation in the model with 15 parameter OLS models where each feature is predicted by its on lags and lags of the other 14 features in the model. The density distribution plots below shows the residuals from the fitted VAR(5) model.

The error terms i.e. the residuals are expected to be independently distributed across each equation and serially uncorrelated. The residuals are experimental errors derived by finding the difference

between the observed data points and the predicted data points. To ensure that these assumptions are met, the researcher tests the hypothesis of white noise residuals using a normality test on the residuals of the VAR model. The stationarity and normality test results for the residuals are shown below.

	Stationarity Test			Normality Test	
	P-value	**Test Statistic**	**Stationary**	**Skewness**	**Kurtosis**
Weaponization	0.01	-10.58	True	0.38	0.78
Joy	0.01	-11.35	True	0.18	0.23
Mean Bytes	0.01	-13.66	True	3.06	42.69
Flu Relatedness	0.01	-12.95	True	0.28	0.54
Sadness	0.01	-12.52	True	0.51	0.69
Fire Relatedness	0.01	-10.58	True	1.47	5.97
Entropy	0.01	-13.56	True	1.06	3.95
Cyber Relatedness	0.01	-11.33	True	0.56	1.12
Sentiment	0.01	-11.70	True	0.53	-0.88
Trust	0.01	-11.81	True	0.44	1.13
Disgust	0.01	-13.11	True	1.61	4.44
Fear	0.01	-10.59	True	0.55	0.50
Anticipation	0.01	-11.55	True	0.64	0.59
Surprise	0.01	-11.46	True	0.61	1.57
Decoy	0.01	-11.11	True	9.66	244.30

Table 5-21: Residual Stationarity and Normality Test Results (Weaponization Phase)

The distribution of the residuals for intrusion detection features are seen to be centered on zero with narrow peaked curves. The model performs moderately in capturing the trends in the other features as the residuals are seen to be slightly skewed to the left.

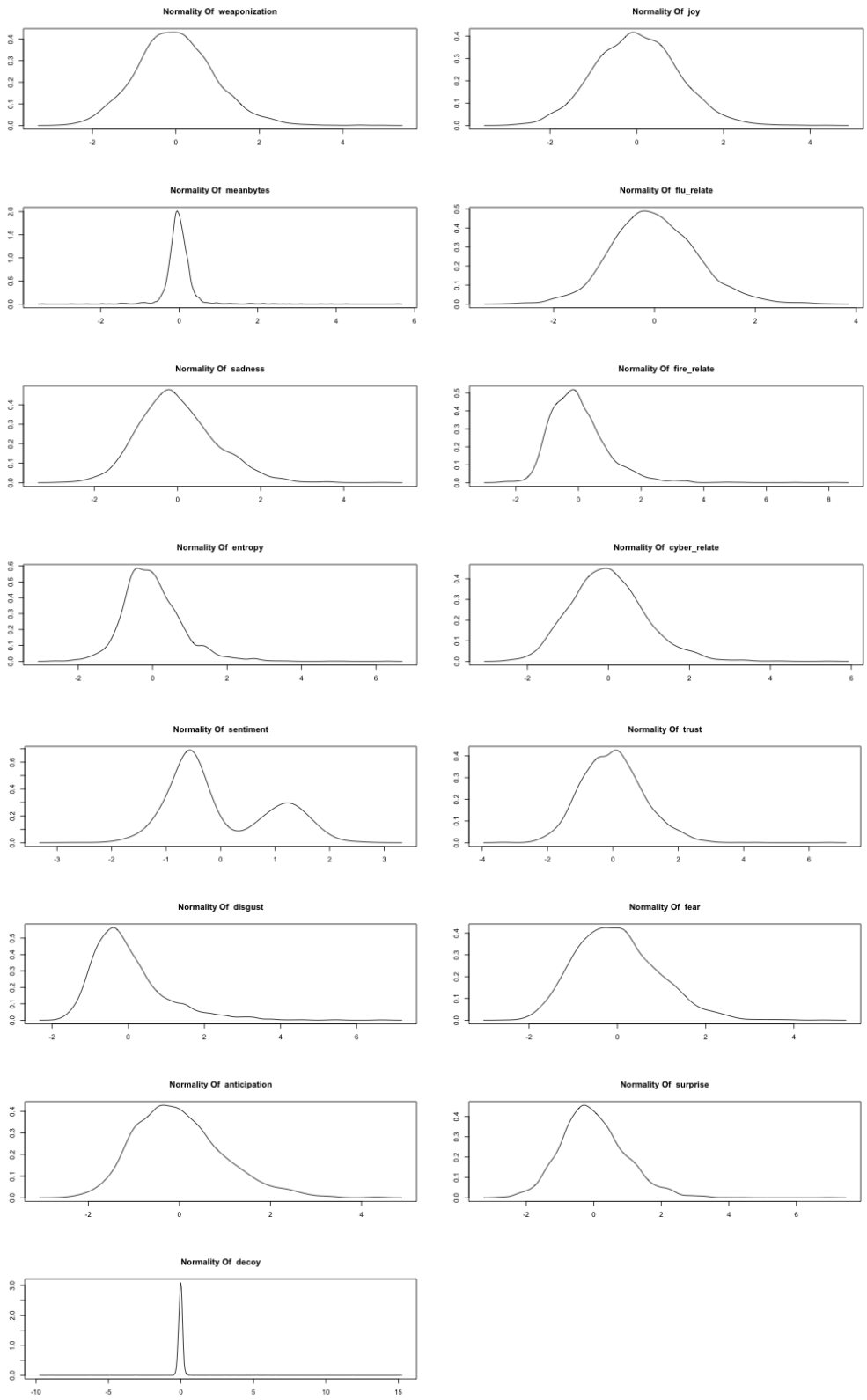

Figure 5-27: Density Distribution of Residuals (Weaponization Phase)

Lastly, after observing the density distribution of the residuals, the researcher derives a normal probability plot for each residual from the structural model. To achieve this, the researcher calculates the cumulative probability of each residual using the formula:

$$P(i-th\ residual) = \frac{i}{(N+1)}$$

Where P is the cumulative probability of an observed residual, I is residual observation and N is the number of observations made within the given time period.

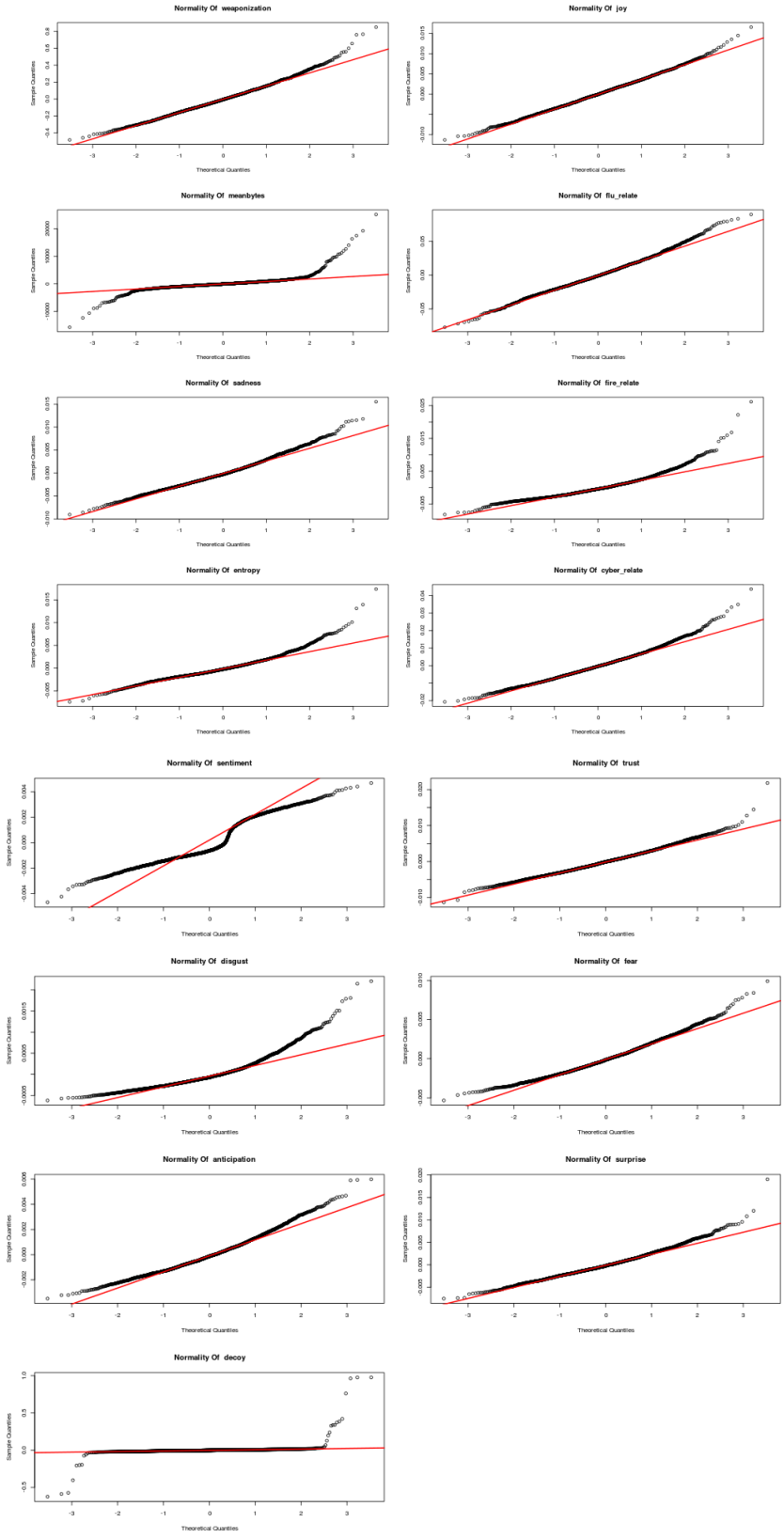

Figure 5-28: Normality Probability Plot of Residuals (Weaponization Phase)

The normal probability plot identifies departures from normality, non-linearity and outliers in a given distribution. Deviations from the straight line indicate a deviation from normality. A homoscedastic

distribution is observed across all plots with reducing errors around zero. The S-Shape formed by data points in the plots indicate a shift towards normality. However, a few outliers are observed in all 15 features.

The performance of the VAR model is measured using structural multivariate time series model validation techniques outlined in the literature review. The model is validation using both in sample and out of sample prediction values. The in-sample residuals are derived by obtaining the differences between the actual observations and the predictions returned by the model. The results for the in-sample model performance is shown in the table below.

	MAE	RMSE	MPE	MAPE	MAD	FEATURE ACCURACY
WEAPONIZATION	0.74	0.94	-891.59	891.59	0.00	66.53
JOY	0.76	0.95	117.34	117.34	0.00	18.85
MEANBYTES	0.21	0.39	18.75	18.75	0.00	80.48
FLU_RELATE	0.66	0.85	-336.31	336.31	0.00	69.83
SADNESS	0.73	0.93	13.60	13.60	0.00	61.51
FIRE_RELATE	0.68	0.92	-107.02	107.02	0.00	68.01
ENTROPY	0.59	0.80	1775.52	1775.52	0.00	-638.10
CYBER_RELATE	0.72	0.93	-42.59	42.59	0.00	67.01
SENTIMENT	0.83	0.96	-1457.74	1457.74	0.00	64.17
TRUST	0.75	0.95	-373.60	373.60	0.00	66.04
DISGUST	0.72	0.98	-419.57	419.57	0.00	66.01
FEAR	0.74	0.93	-61.37	61.37	0.00	66.76
ANTICIPATION	0.76	0.97	754.14	754.14	0.00	-236.35
SURPRISE	0.75	0.96	-219.11	219.11	0.00	65.92
DECOY	0.16	0.77	-384.32	384.32	0.00	81.37
AVEARGE TOTAL ERROR	0.65	0.88	-107.59	464.84	0.00	

Table 5-22: In Sample Model Performance Results (Weaponization Phase)

The model works well in predicting most of the features in the model except the UDP connections feature which records a 637% average error. The model also performed accurately for features with smaller values of RMSPE, MAD and MSPE. These results show that the model works well with expected data and fits the observations of interest.

Additionally, out of sample validations are also done for each feature Out of sample forecasts are obtained by forecasting the N-ahead data values where N=60 minutes i.e the last hour initially subtracted from the original dataset. The out of sample residuals are then derived by obtaining the differences between forecasts on the test datasets using the VAR model and the test data with 60 observations. The measures of out of sample performance values are presented in the table below.

	MAE	RMSE	MPE	MAPE	MAD	FEATURE ACCURACY
WEAPONIZATION	0.74	0.94	-891.59	891.59	0.00	64.4147375
JOY	0.76	0.95	117.34	117.34	0.00	-165.25613

MEANBYTES	0.21	0.39	18.75	18.75	0.00	62.5591361
FLU_RELATE	0.66	0.85	-336.31	336.31	0.00	57.4980546
SADNESS	0.73	0.93	13.60	13.60	0.00	49.9239341
FIRE_RELATE	0.68	0.92	-107.02	107.02	0.00	58.3512524
ENTROPY	0.59	0.80	1775.52	1775.52	0.00	57.7179547
CYBER_RELATE	0.72	0.93	-42.59	42.59	0.00	57.5131851
SENTIMENT	0.83	0.96	-1457.74	1457.74	0.00	58.8138437
TRUST	0.75	0.95	-373.60	373.60	0.00	-24.772427
DISGUST	0.72	0.98	-419.57	419.57	0.00	58.7325357
FEAR	0.74	0.93	-61.37	61.37	0.00	5.64543808
ANTICIPATION	0.76	0.97	754.14	754.14	0.00	56.6288404
SURPRISE	0.75	0.96	-219.11	219.11	0.00	57.1853207
DECOY	0.16	0.77	-384.32	384.32	0.00	-309.58694
AVEARGE TOTAL ERROR	0.65	0.88	-107.59	464.84	0.00	

Table 5-23: Out of Sample Model Performance Results (Weaponization Phase)

The out of sample performance shows that the model works well in predicting new values of the endogenous feature. However, some features such as predictions for the mean packet bytes and the level of cyber relatedness of microblogging feeds are seen to exhibit a higher level of out of sample percentage error. Although these features are selected as being important for predicting the outcome variable, the features included in the model are not entirely efficient in predicting them. By averaging the performance measure across the various performance measurement techniques, the model records a total of 85% in sample accuracy and 67% out of sample accuracy for the outcome feature.

Following the performance test, the granger causality tests each feature combination with the optimum lag selected in the model order selection stage. The structural model can also be used to make inference about the direction or directions of causality between every pair of features in the feature set. The granger causality test is set up with the null hypothesis of no causality between the feature 'x' (on the left-hand of the arrow) and the feature 'y' (on the right-hand of the arrow). The table below shows only causal links with significant P-values to reject the null hypothesis.

	Causal Relations	**F Statistic**	**P-value**
1.	**Weaponization --> Joy**	7.05	0.01
2.	**Weaponization --> Flu Relatedness**	26.11	0.00
3.	**Weaponization --> Sadness**	6.58	0.01
4.	**Weaponization --> Fire Relatedness**	8.33	0.00
5.	**Weaponization --> Entropy**	30.53	0.00
6.	**Joy --> Mean transmitted bytes**	31.33	0.00
7.	**Joy --> Cyber Relatedness**	12.18	0.00
8.	**Joy --> Fear**	10.74	0.00

9.	**Joy --> Anticipation**	5.37	0.02
10.	**Joy --> Decoy**	8.90	0.00
11.	**Mean transmitted bytes --> Joy**	6.14	0.01
12.	**Flu Relatedness --> Weaponization**	10.92	0.00
13.	**Flu Relatedness --> Joy**	8.02	0.00
14.	**Flu Relatedness --> Mean transmitted bytes**	12.75	0.00
15.	**Flu Relatedness --> Entropy**	13.32	0.00
16.	**Flu Relatedness --> Cyber Relatedness**	5.74	0.02
17.	**Flu Relatedness --> Fear**	10.15	0.00
18.	**Flu Relatedness --> Anticipation**	6.15	0.01
19.	**Sadness --> Weaponization**	13.38	0.00
20.	**Sadness --> Joy**	5.25	0.02
21.	**Sadness --> Flu Relatedness**	89.77	0.00
22.	**Sadness --> Entropy**	6.87	0.01
23.	**Sadness --> Cyber Relatedness**	21.74	0.00
24.	**Fire Relatedness --> Flu Relatedness**	7.42	0.01
25.	**Fire Relatedness --> Entropy**	20.91	0.00
26.	**Fire Relatedness --> Trust**	8.34	0.00
27.	**Fire Relatedness --> Surprise**	5.17	0.02
28.	**Entropy --> Weaponization**	12.71	0.00
29.	**Entropy --> Flu Relatedness**	7.95	0.00
30.	**Entropy --> Sadness**	20.05	0.00
31.	**Entropy --> Sentiment**	8.60	0.00
32.	**Entropy --> Trust**	28.58	0.00
33.	**Entropy --> Fear**	34.26	0.00
34.	**Entropy --> Surprise**	22.35	0.00
35.	**Cyber Relatedness --> Flu Relatedness**	22.38	0.00

36.	**Cyber Relatedness --> Sadness**	11.15	0.00
37.	**Cyber Relatedness --> Sentiment**	6.96	0.01
38.	**Cyber Relatedness --> Trust**	10.97	0.00
39.	**Cyber Relatedness --> Fear**	16.51	0.00
40.	**Cyber Relatedness --> Anticipation**	9.24	0.00
41.	**Cyber Relatedness --> Surprise**	9.25	0.00
42.	**Sentiment --> Weaponization**	5.96	0.01
43.	**Sentiment --> Flu Relatedness**	10.81	0.00
44.	**Sentiment --> Cyber Relatedness**	8.37	0.00
45.	**Trust --> Mean transmitted bytes**	5.83	0.02
46.	**Trust --> Entropy**	19.65	0.00
47.	**Trust --> Cyber Relatedness**	11.96	0.00
48.	**Disgust --> Cyber Relatedness**	6.78	0.01
49.	**Fear --> Joy**	10.25	0.00
50.	**Fear --> Mean transmitted bytes**	30.72	0.00
51.	**Fear --> Sadness**	9.36	0.00
52.	**Fear --> Entropy**	23.32	0.00
53.	**Fear --> Cyber Relatedness**	22.60	0.00
54.	**Fear --> Anticipation**	9.47	0.00
55.	**Anticipation --> Mean transmitted bytes**	7.41	0.01
56.	**Surprise --> Flu Relatedness**	6.47	0.01
57.	**Surprise --> Fire Relatedness**	11.59	0.00
58.	**Surprise --> Entropy**	25.75	0.00
59.	**Surprise --> Cyber Relatedness**	7.72	0.01

Table 5-24: F-Test Results for Granger Analysis (Weaponization Phase)

The granger analysis shows a total of 59 likely causal relations of 225 possible relations between the 15 features in the feature set. The weaponization feature is seen to have plausible causal relations with 4 emotion features on the social dimension. There is also an observed bi-causal relation or feedback loop between the weaponization feature and these 4 features. Inter-dimensionally, there exists only relations between features on the social dimension. Intra-dimensionally, the number of transmitted

bytes feature from on the network layer is seen to be linked with emotion features on the social dimension such as fear, trust and anticipation.

5.4.4 Stage 4: The Delivery Stage

Stage 4 of the experiment integrates evidences leading from the physical dimension and evidences on the social dimension on a pre-defined time scale to spot indicators of a plausible injection in victims' network on the physical dimension of cyberspace. The rationale for conducting this stage of the experiment is to establish entanglements between events on the physical and social dimension of cyberspace using co-integral links between them at the 3rd stage of the traditional kill chain model. The researcher uses series from the network layer of the physical dimension considering prior stages of the kill-chain already identified in previous stages and the social dimension in constructing a predictive model.

The endogenous variable selected for this phase characterizes an injection using database and http connections to the target network. The 'injection' variable is derived as a function of identified malicious HTTP connections and Database connections to target network (Roy, Singh and Sairam, 2011; Issues, 2012; Raghava, Sahgal and Chandna, 2012). It is estimated by scaling the number of database connections by the number of malicious HTTP connections, all scaled by the total number of connections made and packets sent over the network.

$$\frac{1}{Total\ Number\ of\ Packets} * \left(\frac{Database\ Requests}{Malicious\ HTTP\ Connections}\right)$$

Equation 5-17: Characterizing an Injection in Network Traffic

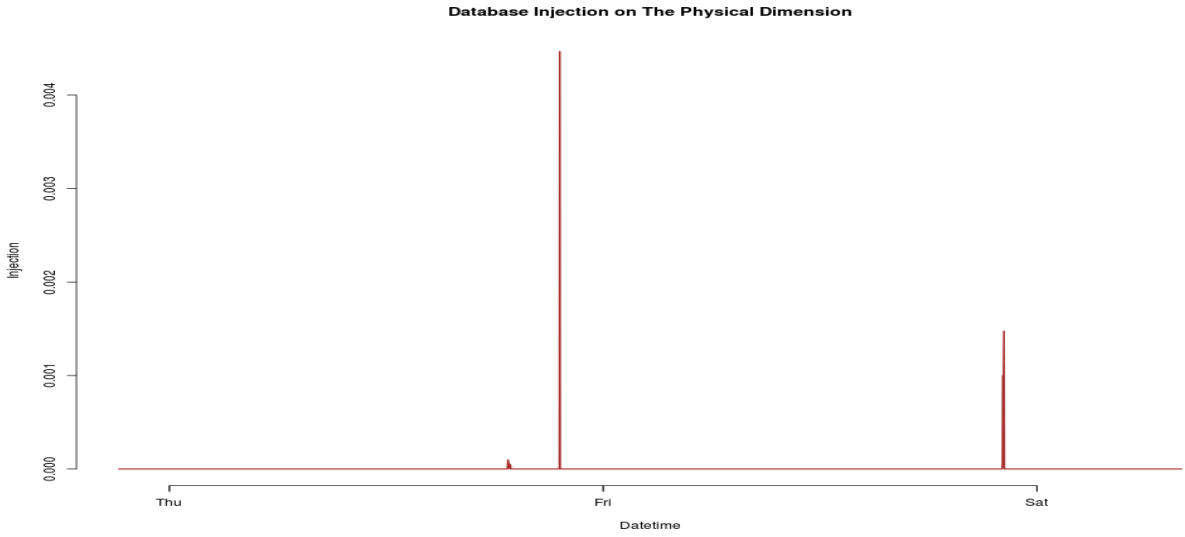

Figure 5-29: Database Injection

The endogenous variable selected at this stage is the 'injection' which indicates an attempt at a database injection on target network, is arbitrary to represent the delivery, exploitation and installation stages of the traditional kill-chain.

Datasets from both dimensions were integrated on a similar timeline with the aim of predicting events of the last hour on the physical dimension. The merging of both datasets produced a single dataset

with 72 features and 2524 observations. For testing and evaluation of developed model, the data was split into two removing the last 120 observations (the two hours) from each feature in the dataset. The out of sample reliability of the constructed model will be tested on observations from the last hour using an N-ahead prediction strategy where N=120. This is done to ensure reliability of out of sample prediction performance testing.

A simple intra-dimensional correlation analysis of the 72 features on both dimensions as shown in the figure below, reveals a slight correlation between features on both dimensions. For example, the level of chat congestion and chatroom traffic is seen to have a slight correlation with the number of distinct connections to the network. Additionally, on the social dimension, emotion and opinion features are observed to be highly correlated with each other. The black boxes in the figure below groups features into hierarchical clusters. The recon, connection fail ratio, scanner and weaponization features are propagated from the reconnaissance and weaponization phases of the kill-chain. Sets of network and social features are observed to be highly correlated with each other such as ip addresses, ports, number of connections denied by intrusion detection system, total alerts raised by intrusion detection system, entropy, weaponization, chat room congestion. These correlation relationships were also observed in the previous stages of the experiment.

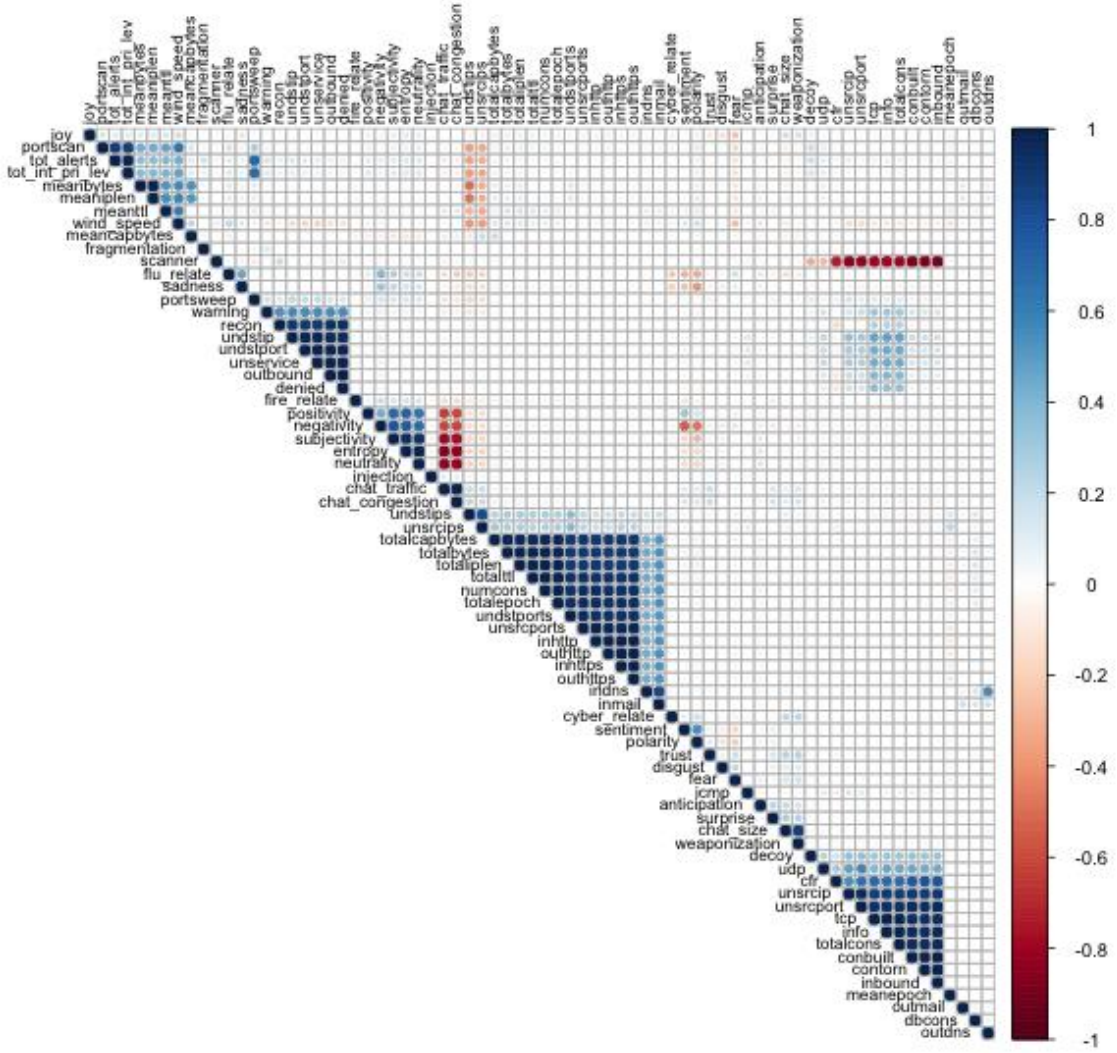

Figure 5-30: Intra-Dimensional Correlation Coefficients (Delivery Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment. The correlation coefficients are on a scale of +1 to -1 as shown by the color bar. Highly positively correlated pairs of variables tend towards +1 while highly negatively correlated pairs of variables tend towards -1. Multiple pairs of variables are seen to be highly positively or negatively correlated. For example, the group of features, Cyber relatedness ('cyber_relate'), Entropy, chat size ('chat_size'), chat room congestion ('chat_congestion') and chat room traffic ('chat_traffic') posted per minute are all seen to be highly positively correlated. Additionally, the group of features such as packet data features, reconnaissance features and intrusion detection features are also seen to be highly correlated with each other. Consequently, one or more of pairs of features that are seen to be highly correlated with each other (where Pearson's correlation coefficient $\geq +0.65$ or ≤ -0.65) are removed from the dataset. The correlation elimination process reduces the feature set by removing one or more of a set of highly correlated variables (Hall and Smith, 1998). The assumption is that information provided by highly correlated variables can be provided to the model by a single one of those variables therefore reducing redundancy to achieve a parsimonious model. The results of the filtered feature set after a correlation

elimination produces a smaller feature set of 24 features as shown in the figure below, eliminating 48 highly correlated features.

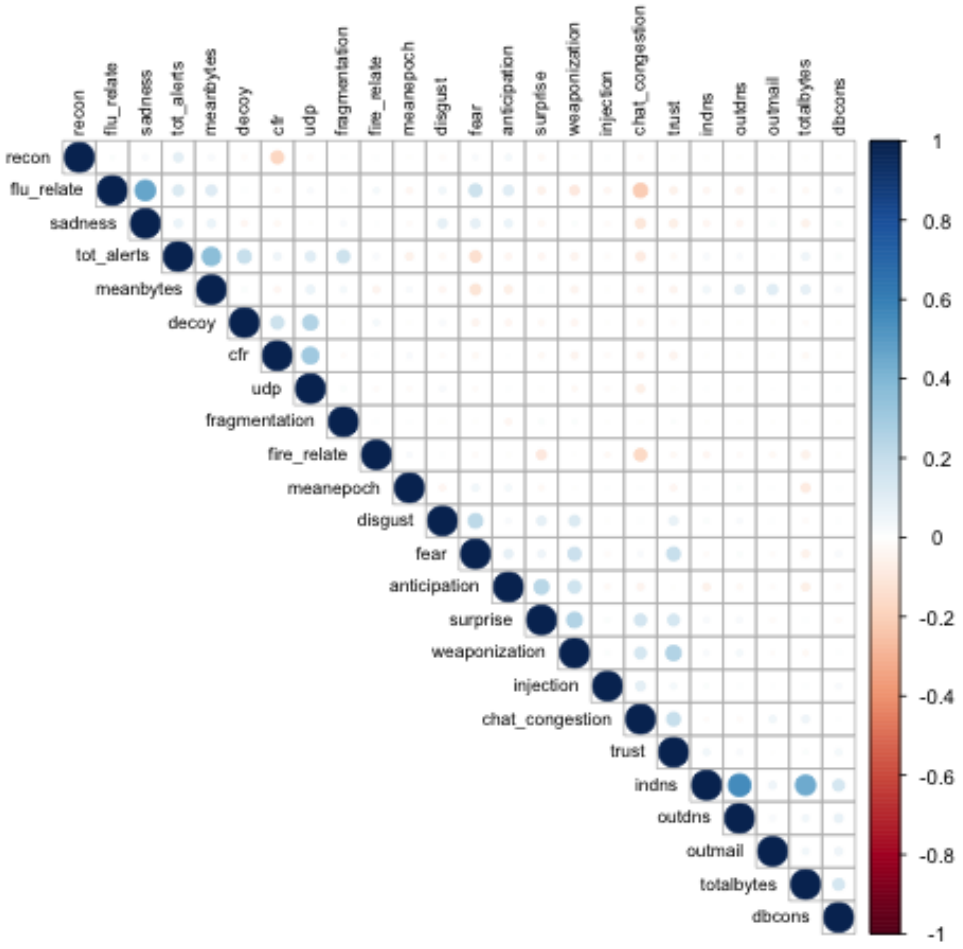

Figure 5-31: Correlation Co-efficient of Reduced Feature set (Delivery Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment for the reduced feature set relevant to the delivery stage of the kill-chain. This reduced set of 24 features including the endogenous variable, 'injection' is selected for further feature selection methods. To further reduce the number of features, the researcher applies a recursive features selection method to select only features that significantly reduce the error of prediction in the endogenous feature. The recursive feature selection stage further reduces the uncorrelated feature set to only those variables that minimizes the prediction error for the chosen endogenous variable 'Injection'. To begin the recursive feature selection from the uncorrelated feature set, the researcher starts by building an ordinary least squares regression linear model with the 'injection' as the endogenous feature predicted by all other features in the reduced feature set above as shown in the figure above. The recursive feature selection searches for a model that optimizes the Akaike Information (AIC) and reduces errors. The recursive feature selection step further reduces the number of features leaving 16 features. The resulting features are shown in the table below and used in further analysis. Lastly, the researcher creates generic methods for calculating the importance of each feature in predicting the outcome feature. This is used to examine the contribution of each feature in the dataset in predicting the outcome feature 'injection'. This last step in the feature selection phase is elimination by variable importance (Mehmood *et al.*, 2012). The absolute value of the t-statistic for each feature is used to compute the

unique contribution of each input-feature to the model and a cut off of 1 is chosen (Noppamas, Seree and Kidakan, 2014). From the figure below, it is observed that the most important features for predicting the outcome feature 'injection' are the total number of database connections, entropy/chatroom congestion and population fear and trust derived from microblogging feeds.

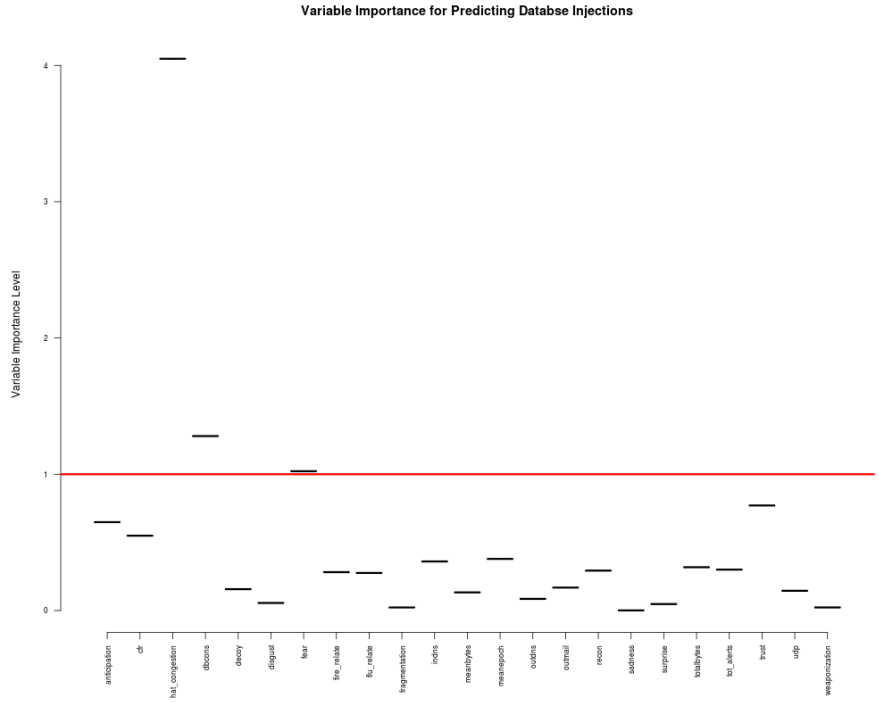

Figure 5-32: Variable Importance Of Features (Delivery Phase)

The figure above shows the relative variable importance for each feature in the final feature set. The last phase of the feature selection left the same homogenous features as extremely useful for predicting the outcome feature. The final features for the model are shown in table 5-24 below.

The next phase of the analytical framework involves selecting appropriate orders for the intended model. This involves selecting an autoregressive order, testing for seasonality in each feature and selecting a moving average as discussed in the analytical framework.

The VAR model order selection phase seeks to select the appropriate significant lag at which features are serially correlated. The researcher uses an iterative solution with the maximum of 100 lags. At each iteration, the researcher builds a model for each equation in the VAR and measures the AIC. Finally, the researcher selects 15 lags (approximately 15 minutes) as this is the lag length with the optimal AIC.

In addition to selecting appropriate VAR lag order, the researcher also tests each time series in the feature set for seasonality. In order to achieve this, the researcher employs a time series decomposition technique. As discussed in the literature review, time series decomposition works by splitting a time series into its three main components: the trend, the seasonal movement and the error terms. Seasonal patterns are expected to repeat with a fixed period of time. An additive decomposition method was chosen as the magnitude of the series does not increase with the series. After applying the Fourier transform for detecting seasonality, the injection feature shows an approximately 2-hour seasonal variation while the database connections show an approximately 2.5 hour seasonal variation.

A stationarity and normality test are therefore conducted on the 4 features and results are tabled below. The results of the stationarity test show a significant p-value for rejecting the null hypothesis

of non-stationarity in the augmented dickey fuller stationarity test. All selected features are seen to have a constant mean and variance over the selected period of time.

	Stationarity Test			Normality Test			
	P-Value	**Augmented Dickey Fuller Statistic**	**Stationary**	**Skewness**	**Kurtosis**	**Kolmogorov–Smirnov Test Statistic**	**K-S P-Value**
Injection	0.01	-12.57	True	36.13	1483.21	0.50	0.00
Chat Congestion	0.01	-5.42	True	1.97	6.92	1.00	0.00
Fear	0.01	-8.59	True	0.50	0.31	0.50	0.00
Database Connections	0.01	-11.60	True	24.92	644.54	0.50	0.00

Table 5-25: Features' Stationarity and Normality Test Results (Delivery Phase)

Additionally, the skewness and K-S tests show that the Chat Congestion and population fear features are fairly normal. However, the Injection and database connections are observed to be slightly right skewed all features exhibiting varying peakedness shown by the kurtosis test. Given that all features were stationary, the Johansen's test for co-integration was not applicable at this stage of the experiment. However, for further investigation, the Engle and Granger test for co-integration was used to determine co-integration between each pair of feature in the feature set. To achieve this, the researcher constructs a linear regression model between each pair of features and conducts a stationarity test on the residuals of the resulting model. The Engle Granger test shows no significant p-values for features in the system at the selected time period.

Given that all features are observed to be stationary and the absence of any co-integrating relationships at this stage of the experiment, a VAR model was fitted to the levels. The VAR with order of 15 lags was estimated utilizing the OLS per equation in the model with 4 parameter OLS models where each feature is predicted by its on lags and lags of the other 3 features in the model. The density distribution plots below shows the residuals from the fitted VAR(15) model.

The error terms i.e. the residuals are expected to be independently distributed across each equation and serially uncorrelated. The residuals are experimental errors derived by finding the difference between the observed data points and the predicted data points. To ensure that these assumptions are met, the researcher tests the hypothesis of white noise residuals using a normality test on the residuals of the VAR model. The stationarity and normality test results for the residuals are shown below.

	Stationarity Tests			**Normality Test**			
	P-Value	**Augmented Dickey Fuller Test Statistic**	**Stationary**	**Skewness**	**Kurtosis**	**Kolmogorov–Smirnov Test Statistic**	**K-S P-Value**
Injection	0.01	-13.06	1.00	35.42	1493.90	0.38	0.00
Chat Congestion	0.01	-12.99	1.00	0.74	4.47	0.12	0.00
Fear	0.01	-13.55	1.00	0.55	0.58	0.04	0.00
Database Connection	0.01	-13.06	1.00	8.69	438.52	0.44	0.00

s							

Table 5-26: Residuals' Stationarity and Normality Test Results (Delivery Phase)

The distribution of the residuals for the outcome feature injection is seen to be centered at zero with narrow peaked curves and slightly skewed to the right. The distribution curve shows that the model performs well in predicting the database connections and the chat congestion features. The model performs moderately in capturing the trends in the other features as the residuals are seen to be slightly skewed to the right.

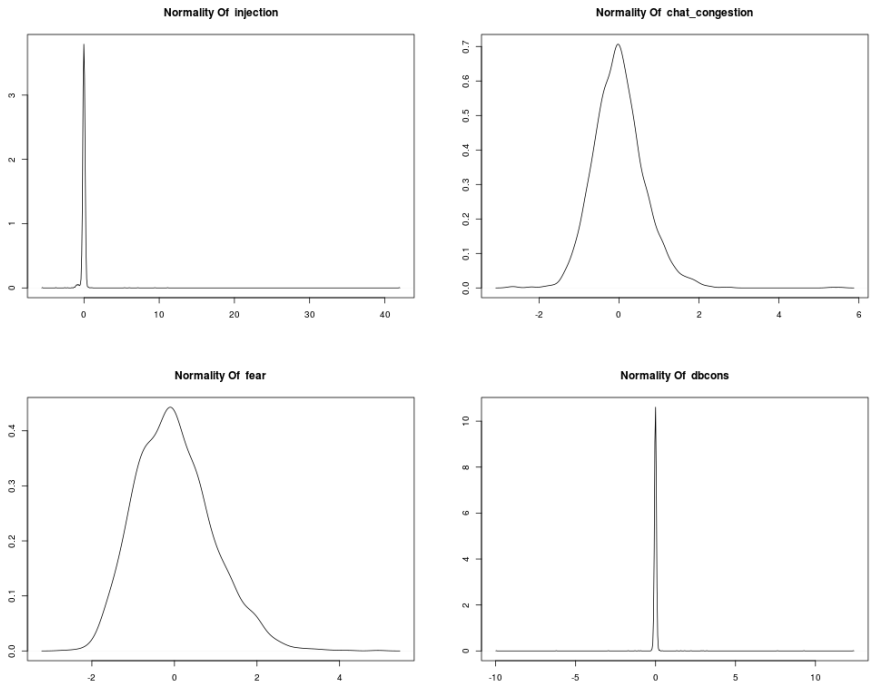

Figure 5-33: Density Distribution of Residuals (Delivery Phase)

Lastly, after observing the density distribution of the residuals, the researcher derives a normal probability plot for each residual from the structural model. To achieve this, the researcher calculates the cumulative probability of each residual using the formula:

$$P(i-th\ residual) = \frac{i}{(N+1)}$$

Where P is the cumulative probability of an observed residual, I is residual observation and N is the number of observations made within the given time period.

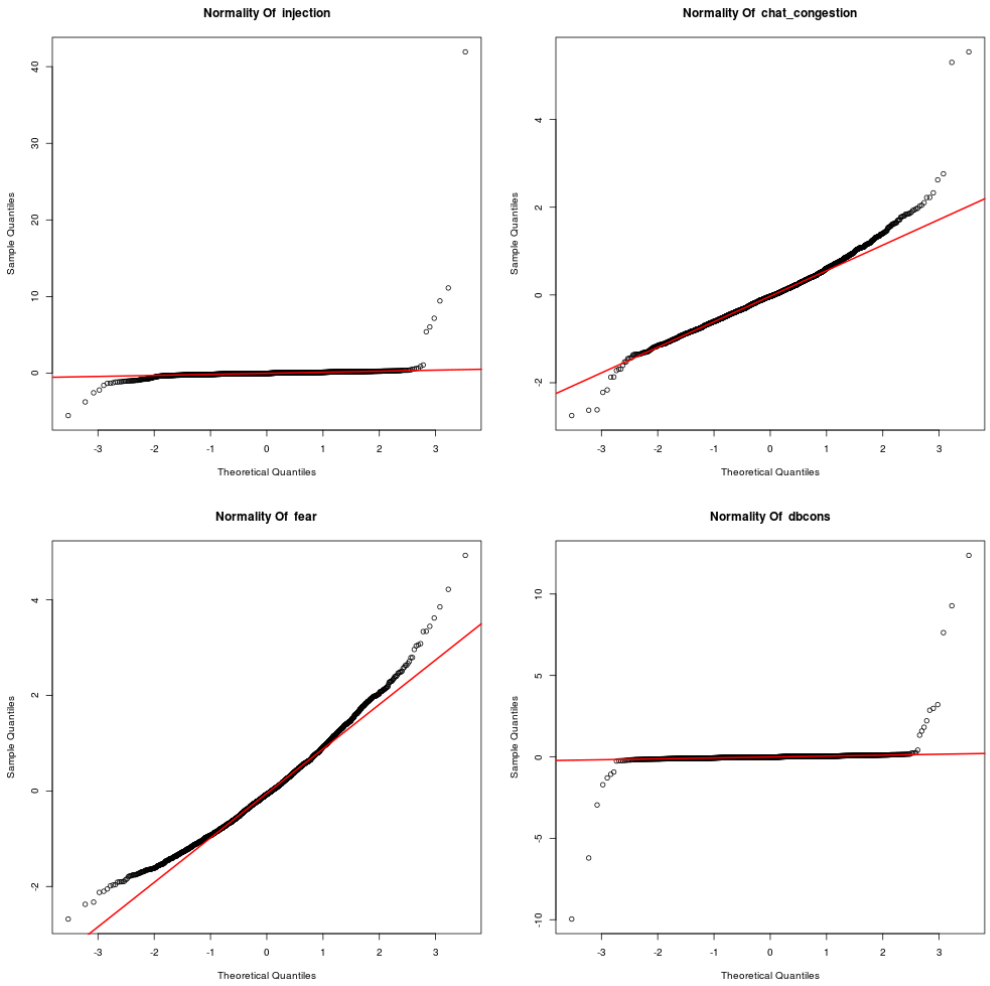

Figure 5-34: Normality Probability Plot of Residuals (Delivery Phase)

The normal probability plot identifies departures from normality, non-linearity and outliers in a given distribution. Deviations from the straight line indicate a deviation from normality. A homoscedastic distribution is observed across all plots with reducing errors around zero. The S-Shape formed by data points in the plots indicate a shift towards normality. However, a few outliers are observed in all 4 features.

The performance of the VAR model is measured using structural multivariate time series model validation techniques outlined in the analytical framework. The model is validation using both in sample and out of sample prediction values. The in-sample residuals are derived by obtaining the differences between the actual observations and the predictions returned by the model. The results for the in-sample model performance is shown in the table below.

	MAE	RMSE	MPE	MAPE	MAD	FEATURE ACCURACY
INJECTION	0.17	0.97	-45.66	45.66	0.00	77.37
CHAT_CONGESTION	0.49	0.65	-33.82	33.82	0.00	77.13
FEAR	0.74	0.93	-813.07	813.07	0.00	66.59
DBCONS	0.07	0.46	-81.65	81.65	0.00	89.37

TOTAL AVERAGE	0.37	0.75	-243.55	243.55	0.00	77.61

Table 5-27: In Sample Model Performance Results (Delivery Phase)

The model works well in predicting most of the features in the model except the population fear feature which records a 197% average error. The model also performed accurately for features with smaller values of RMSE, MAD and MAPE. These results show that the model works well with expected data and fits the observations of interest.

Additionally, out of sample validations are also done for each feature Out of sample forecasts are obtained by forecasting the N-ahead data values where N=120 minutes i.e the last hour initially subtracted from the original dataset. The out of sample residuals are then derived by obtaining the differences between forecasts on the test datasets using the VAR model and the test data with 120 observations. The measures of out of sample performance values are presented in the table below.

	MAE	**RMSE**	**MPE**	**MAPE**	**MAD**	**FEATURE ACCURACY**
INJECTION	0.17	0.97	-45.66	45.66	0.00	96.12
CHAT_CONGESTION	0.49	0.65	-33.82	33.82	0.00	80.04
FEAR	0.74	0.93	-813.07	813.07	0.00	21.89
DBCONS	0.07	0.46	-81.65	81.65	0.00	73.65
TOTAL AVERAGE	0.37	0.75	-243.55	243.55	0.00	77.61

Table 5-28: Out of Sample Model Performance Results (Delivery Phase)

The out of sample performance shows that the model works well in predicting new values of the endogenous feature. However, some features such as predictions for database connections per minute and population fear of microblogging feeds are seen to exhibit a higher level of out of sample percentage error. Although these features are selected as being important for predicting the outcome variable, the features included in the model are not entirely efficient in predicting them. By averaging the performance measure across the various performance measurement techniques, the model records a total of 85% in sample accuracy and 67% out of sample accuracy for the outcome feature.

Following the performance test, the granger causality tests each feature combination with the optimum lag selected in the model order selection stage. The structural model can also be used to make inference about the direction or directions of causality between every pair of feature in the feature set. The granger causality test is set up with the null hypothesis of no causality between the feature ‘x’ (on the left-hand of the arrow) and the feature ‘y’ (on the right-hand of the arrow). The table below shows only causal links with significant P-values to reject the null hypothesis.

	Causal	F-statistics	P-value
1	**Injection --> Chat Congestion**	31.7184	1.99E-08
2	**Chat Congestion --> Fear**	6.783119	0.009259

Table 5-29: F-Test Results for Granger Analysis (Delivery Phase)

The granger analysis shows a total of 2 likely causal relations of 16 possible relations between the 4 features in the feature set. The outcome feature ‘injection’ is seen to have plausible causal relations with 1 feature on the social dimension. There are no observed bi-causal relation or feedback loop between features at this stage of the experiment. Inter-dimensionally, there exists only relations between features on the social dimension such as the population fear and chat congestion. Intra-dimensionally, the injection feature from on the network layer is seen to be linked with the chat congestion on the social dimension.

5.4.5 Stage 5: The Exploitation Phase

Stage 5 of the experiment integrates evidences leading from the physical dimension and evidences on the social dimension on a pre-defined time scale to spot indicators of a plausible detection of botnets in victim's network on the physical dimension of cyberspace. According to (Sans Institute, 2012), the identification of botnets on a victim's network is also a strong indicator of phases 3, 4 and 5 of the traditional cyber-attack kill-chain (i.e exploitation, installation and C&C). The rationale for conducting this stage of the experiment is to establish entanglements between events on the physical and social dimension of cyberspace using co-integral links between them at the 4th stage of the traditional kill chain model. The researcher uses series from the network layer of the physical dimension while propagating prior stages of the kill-chain already identified in previous stages and the social dimension in constructing a predictive model.

The endogenous variable selected for this phase characterizes the presence of a botnet on victim's network as a ratio of the number of DNS requests on port 53 to the number of UDP packets sent on the victims' network (Sans Institute, 2012). The 'botnets' variable is therefore derived as a function of the total number of incoming and outgoing DNS requests and the total amount of UDP packets sent within selected time intervals on the victim's network. It is estimated by scaling the total number of DNS requests by the number of UDP connections, all scaled by the total number of connections made and packets sent over the network.

$$\frac{1}{Total\ Number\ of\ Connections} * \left(\frac{Incoming\ DNS + Outgoing\ DNS}{Total\ UDP\ Packets}\right)$$

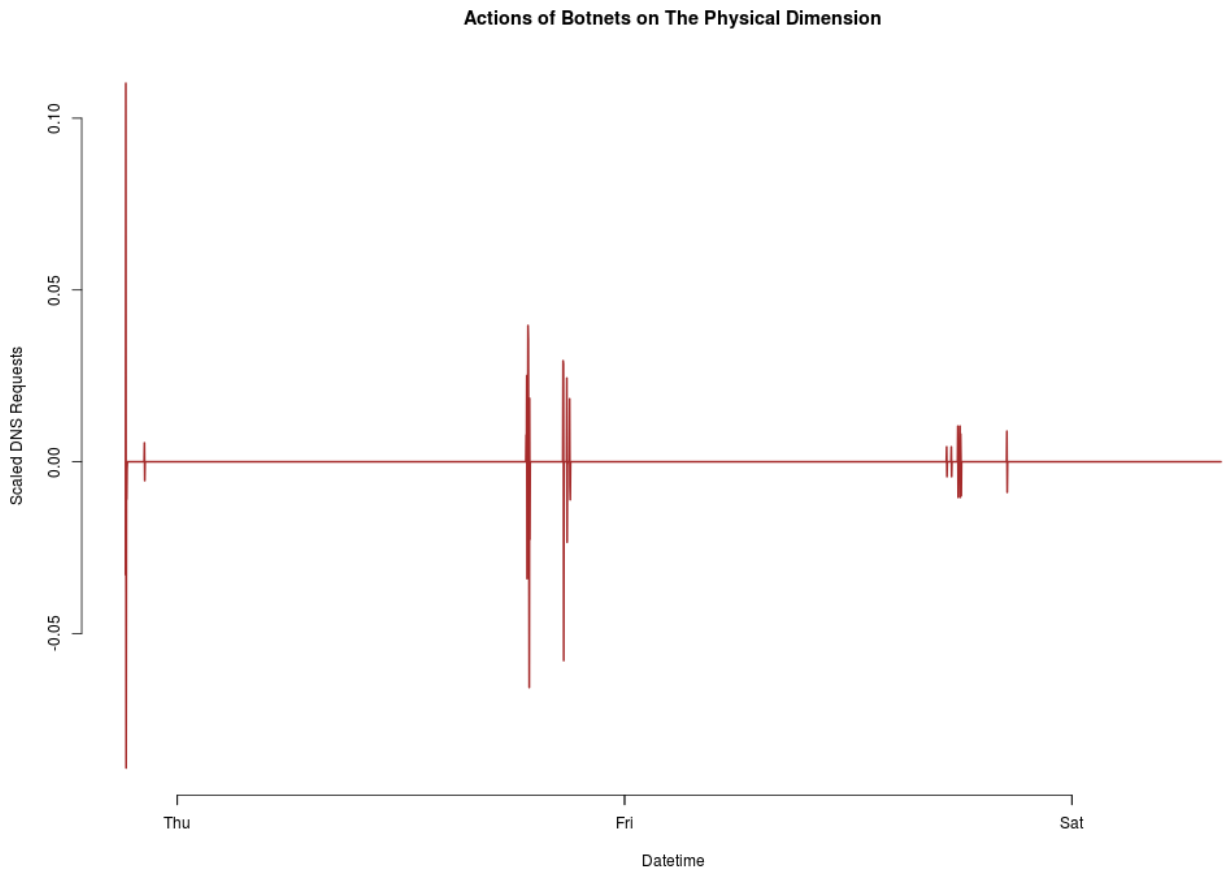

Figure 5-35: Actions of Botnets on the Physical Dimension

The endogenous variable selected at this stage is the 'botnets' which indicates nodes on victim's network are controlled by external adversaries. As with the previous stage of the experiment, this feature is also arbitrary to represent multiple stages of the traditional kill-chain. Datasets from both dimensions were integrated on a similar timeline with the aim of predicting events of the last hour on the physical dimension. The merging of both datasets produced a single dataset with 73 features and 2524 observations. For testing and evaluation of developed model, the data was split into two removing the last 120 observations (the two hours) from each feature in the dataset. The out of sample reliability of the constructed model will be tested on observations from the last hour using an N-ahead prediction strategy where N=120. This is done to ensure reliability of out of sample prediction performance testing.

A simple intra-dimensional correlation analysis of the 73 features on both dimensions as shown in the figure below, reveals a slight correlation between features on both dimensions. For example, the level of chat congestion and chatroom traffic is seen to have a slight correlation with the number of distinct connections to the network. Additionally, on the social dimension, emotion and opinion features are observed to be highly correlated with each other. The black boxes in the figure below groups features into hierarchical clusters. The recon, connection fail ratio, scanner, weaponization and injection features are propagated from the reconnaissance, weaponization and delivery phases of the kill-chain. Sets of network and social features are observed to be highly correlated with each other such as ip addresses, ports, number of connections denied by intrusion detection system, total alerts raised by intrusion detection system, entropy, weaponization, chat room congestion. These correlation relationships were also observed in the previous stages of the experiment.

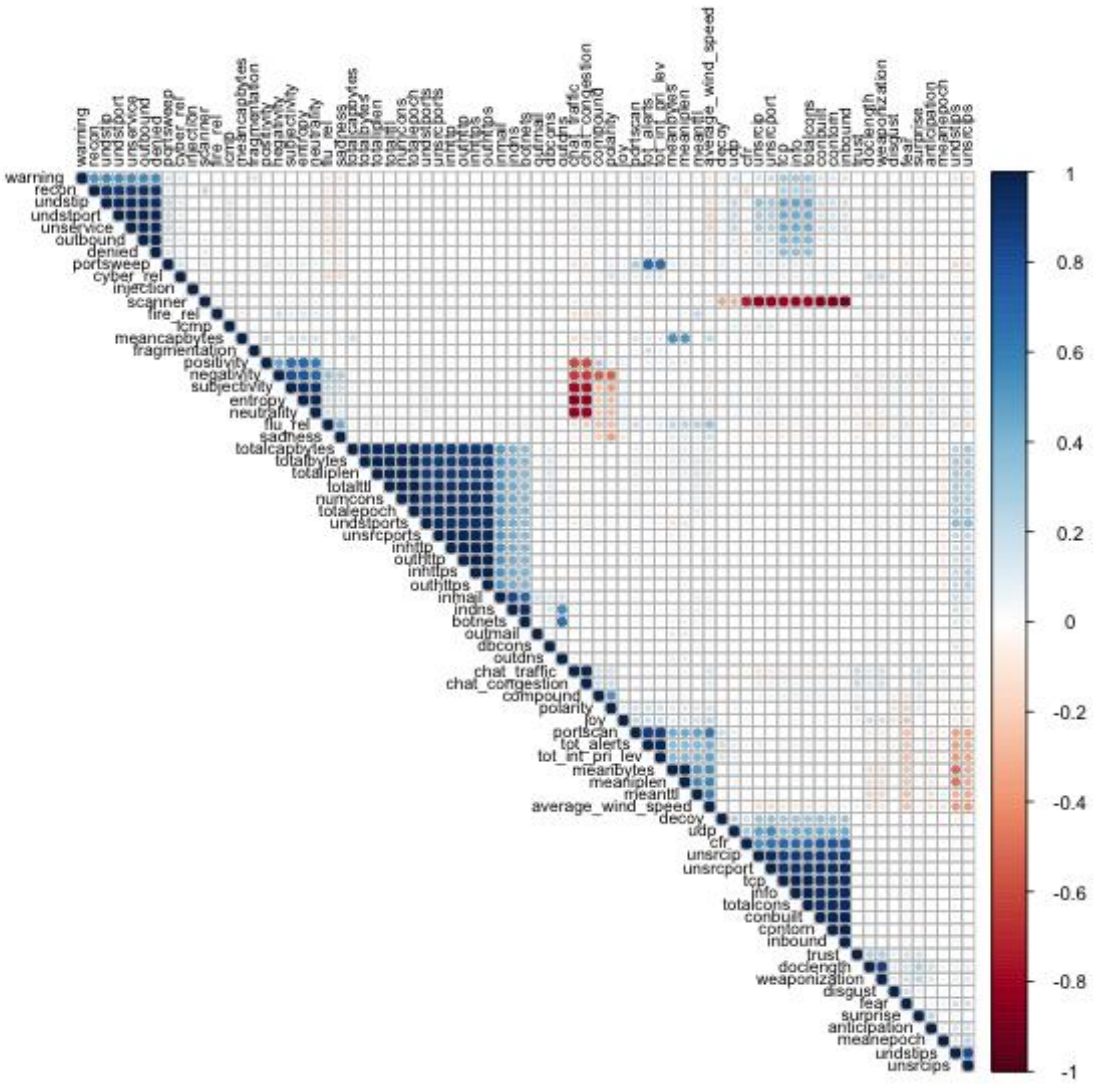

Figure 5-36: Intra-Dimensional Correlation Coefficients (Exploitation Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment relevant to the exploitation phase of the kill-chain. The correlation coefficients are on a scale of +1 to -1 as shown by the color bar. Highly positively correlated pairs of variables tend towards +1 while highly negatively correlated pairs of variables tend towards -1. Multiple pairs of variables are seen to be highly positively or negatively correlated. For example, the group of features, Cyber relatedness ('cyber_relate'), Entropy, chat size ('chat_size'), chat room congestion ('chat_congestion') and chat room traffic ('chat_traffic') posted per minute are all seen to be highly positively correlated. Additionally, the group of features such as packet data features (totalPackets, totalBytes, mean_ip_len, meanttl, undstports, meansrcports), reconnaissance features (portscan, pingsweep, recon, scanner) and intrusion detection features (tot_alerts, warning, decoy) are also seen to be highly correlated with each other. Consequently, one or more of pairs of features that are seen to be highly correlated with each other (where Pearson's correlation coefficient $\geq +0.65$ or ≤ -0.65) are removed from the dataset. The correlation elimination process reduces the feature set by removing one or more of a set of highly correlated variables (Hall and Smith, 1998). The assumption is that information

provided by highly correlated variables can be provided to the model by a single one of those variables therefore reducing redundancy to achieve a parsimonious model. The results of the filtered feature set after a correlation elimination produces a smaller feature set of 27 features as shown in the figure below, eliminating 46 highly correlated features.

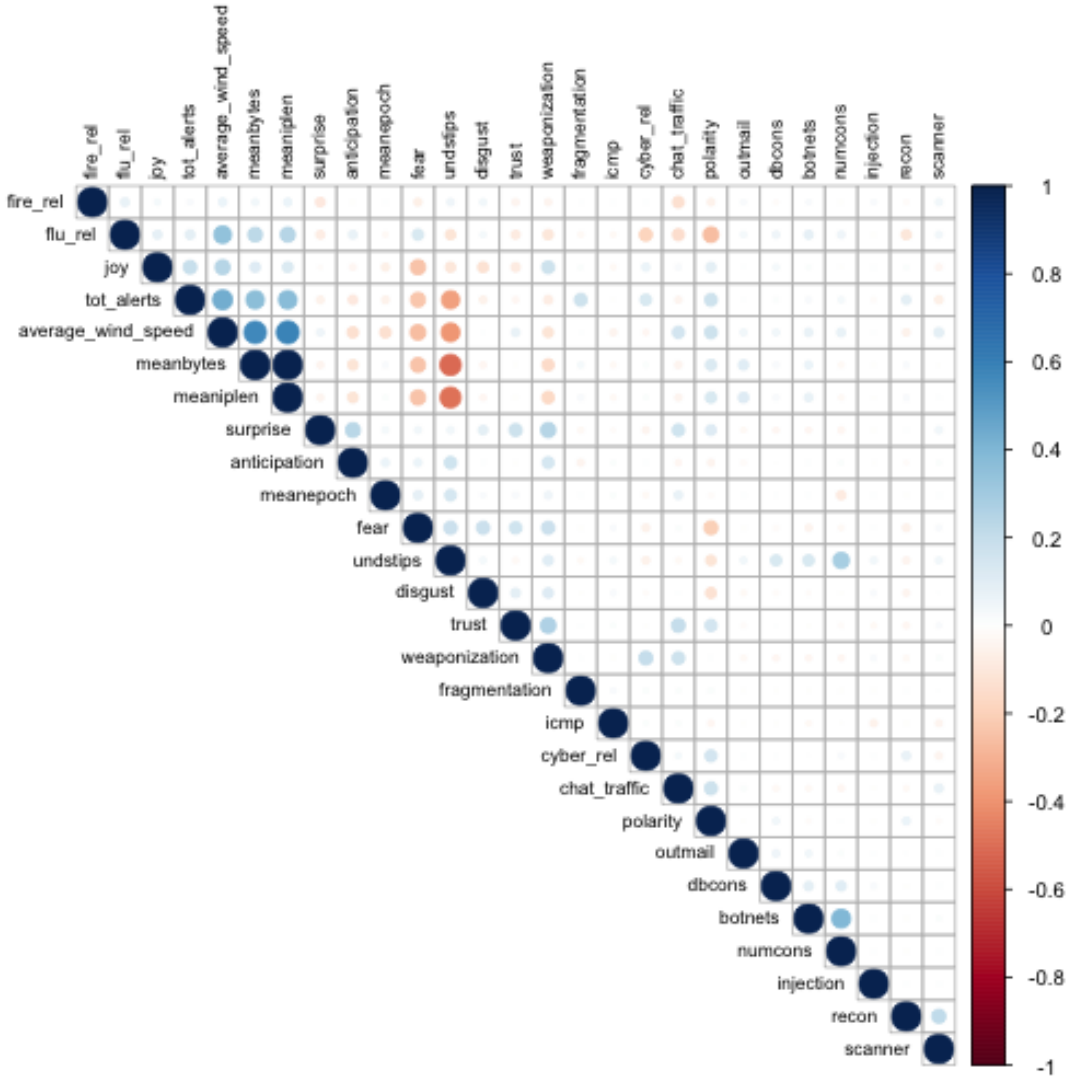

Figure 5-37: Correlation Coefficients Reduced Feature Set (Exploitation Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment relevant to the exploitation phase of the kill-chain for the reduced feature set. This reduced set of 27 features including the endogenous variable, 'botnets' are selected for further feature selection methods. To further reduce the number of features, the researcher applies a recursive features selection method to select only features that significantly reduce the error of prediction in the endogenous feature. The recursive feature selection stage further reduces the uncorrelated feature set to only those variables that minimizes the prediction error for the chosen endogenous variable 'botnets'. To begin the recursive feature selection from the uncorrelated feature set, the researcher starts by building an ordinary least squares regression linear model with the 'botnets' as the endogenous feature predicted by all other features in the reduced feature set above as shown in the figure above. The recursive feature selection searches for a model that optimizes the Akaike Information (AIC) and reduces errors. The recursive feature selection step further reduces the number of features leaving 16 features.

The resulting features are shown in the table below and used in further analysis. Lastly, the researcher creates generic methods for calculating the importance of each feature in predicting the outcome feature. This is used to examine the contribution of each feature in the dataset in predicting the outcome feature 'botnets'. This last step in the feature selection phase is elimination by variable importance (Mehmood *et al.*, 2012). The absolute value of the t-statistic for each feature is used to compute the unique contribution of each input-feature to the model and a cut off of 1 is chosen (Noppamas, Seree and Kidakan, 2014). From the figure below, it is observed that there are 8 important features for predicting the activities of botnets on victim's network. These features include are the total number of database connections, flu-relatedness of micro-blogging feeds, the mean transmitted packet bytes, the total number of packets transmitted on the network and the population fear and trust derived from microblogging feeds.

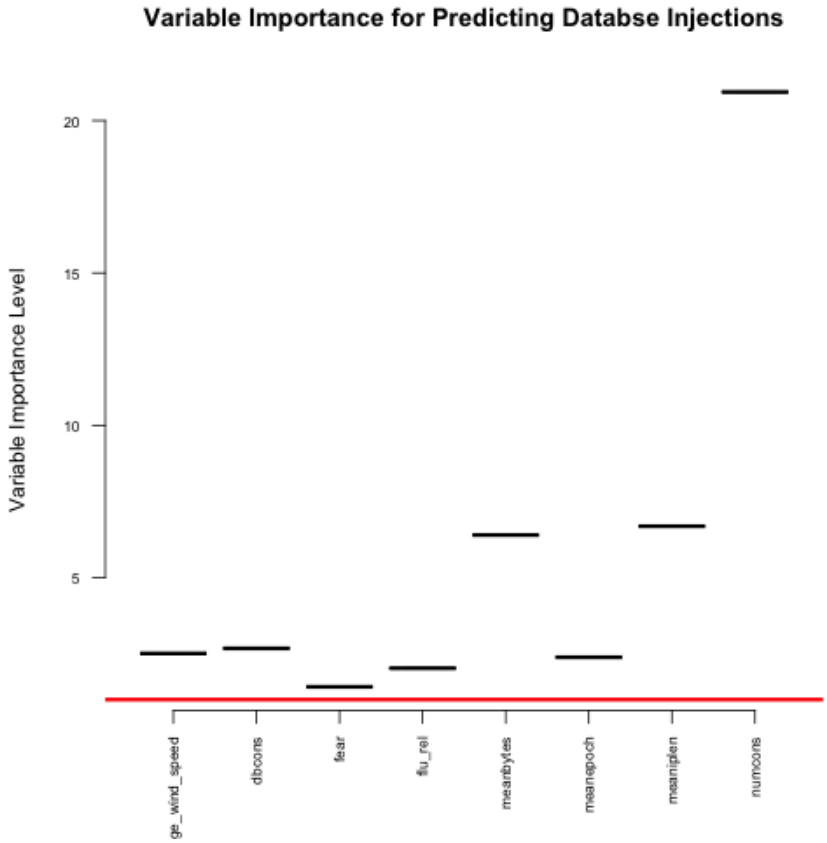

Figure 5-38: Variable Importance of Features (Exploitation Phase)

The figure above shows the relative variable importance for each feature in the final feature set. The last phase of the feature selection left the same 8 homogenous features as extremely useful for predicting the outcome feature. The final features for the model are shown in the table 5-29 below. Features like the average_wind_speed, flu relatedness of microblog feed ('flu_relate') and fear are seen to be important indicators for the exploitation phase. In the context of the scenario, these results reflect the link between the effects of Delish corporation's activities and the population's response to these activities. These 'event-based' features such as tracking the flu relatedness of user discussions ('flu_relate'), given the events in the scenario, are inferred to be useful for pin-pointing adversary activities within the target network.

The next phase of the analytical framework involves selecting appropriate orders for the intended model. This involves selecting an autoregressive order, testing for seasonality in each feature and selecting a moving average as discussed in the analytical framework.

The VAR model order selection phase seeks to select the appropriate significant lag at which features are serially correlated. The researcher uses an iterative solution with the maximum of 300 lags. At each iteration, the researcher builds a model for each equation in the VAR and measures the AIC. Finally, the researcher selects 240 lags (approximately 4 hours) as this is the lag length with the optimal AIC.

In addition to selecting appropriate VAR lag order, the researcher also tests each time series in the feature set for seasonality. In order to achieve this, the researcher employs a time series decomposition technique. As discussed in section literature review, time series decomposition works by splitting a time series into its three main components: the trend, the seasonal movement and the error terms. Seasonal patterns are expected to repeat with a fixed period of time. An additive decomposition method was chosen as the magnitude of the series does not increase with the series. After applying the Fourier transform for detecting seasonality, the botnets and database connections features shows an approximately 2.5 hours seasonal variation.

A stationarity and normality test is therefore conducted on the 9 features and results are tabled below. The results of the stationarity test show a significant p-value for rejecting the null hypothesis of non-stationarity in the augmented dickey fuller stationarity test. All selected features are seen to have a constant mean and variance over the selected period of time.

	Stationarity Test			Normality Test			
	P-Value	**Augmented Dickey Fuller Statistic**	**Stationary**	**Skewness**	**Kurtosis**	**Kolmogorov–Smirnov Test Statistic**	**K-S P-Value**
Botnets	0.01	-11.60	True	18.35	382.97	0.50	0.00
Flu Relatedness	0.01	-8.26	True	0.40	0.48	0.53	0.00
Total Number Of Transmitted Packets	0.01	-12.02	True	18.24	361.81	1.00	0.00
Database Connections	0.01	-11.60	True	24.92	644.54	0.50	0.00
Mean Transmitted Bytes	0.01	-7.29	True	1.92	8.49	1.00	0.00
Mean Ip Length	0.01	-7.25	True	1.82	7.64	1.00	0.00
Average Wind Speed	0.63	-1.88	True	0.20	-1.96	1.00	0.00
Fear	0.01	-10.27	True	0.62	0.77	0.50	0.00
Mean Epoch	0.01	-16.43	True	-16.84	323.15	1.00	0.00

Table 5-30: Features' Stationarity and Normality Test Results (Exploitation Phase)

Additionally, the skewness and K-S tests show that the botnets, total number of transmitted packets, database connections and mean epoch are slightly skewed and exhibiting varying peakedness. However, the kurtosis tests for the flu relatedness, average wind speed and fear are seen to be fairly

normal. Given that all features were stationary, the Johansen's test for co-integration was not applicable at this stage of the experiment. However, for further investigation, the Engle and Granger test for co-integration was used to determine co-integration between each pair of features in the feature set. To achieve this, the researcher constructs a linear regression model between each pair of features and conducts a stationarity test on the residuals of the resulting model. The Engle Granger test shows no significant p-values for features in the system at the selected time period.

Given that all features are observed to be stationary and the absence of any co-integrating relationships at this stage of the experiment, a VAR model was fitted to the levels. The VAR with order of 240 lags was estimated utilizing the OLS per equation in the model with 9 parameter OLS models where each feature is predicted by its on lags and lags of the other 8 features in the model. The density distribution plots below shows the residuals from the fitted VAR(240) model.

The error terms i.e. the residuals are expected to be independently distributed across each equation and serially uncorrelated. The residuals are experimental errors derived by finding the difference between the observed data points and the predicted data points. To ensure that these assumptions are met, the researcher tests the hypothesis of white noise residuals using a normality test on the residuals of the VAR model. The stationarity and normality test results for the residuals are shown below.

	Stationarity Test			Normality Test			
	P-Value	**Augmented Dickey Fuller Statistic**	**Stationary**	**Skewness**	**Kurtosis**	**Kolmogorov–Smirnov Test Statistic**	**K-S P-Value**
Botnets	0.01	-21.93	1.00	0.12	3.70	0.50	0.00
Flu Relatedness	0.01	-20.26	1.00	-0.01	4.01	0.47	0.00
Total Number Of Transmitted Packets	0.01	-18.98	1.00	0.08	4.57	0.50	0.00
Database Connections	0.01	-19.07	1.00	-0.08	4.56	0.49	0.00
Mean Transmitted Bytes	0.01	-20.29	1.00	-0.11	3.93	0.49	0.00
Mean Ip Length	0.01	-20.22	1.00	-0.12	3.92	0.49	0.00
Average Wind Speed	0.01	-18.87	1.00	-0.09	4.54	0.50	0.00
Fear	0.01	-20.17	1.00	0.01	4.01	0.46	0.00
Mean Epoch	0.01	-21.28	1.00	-0.04	4.05	0.49	0.00

Table 5-31: Residuals' Stationarity and Normality Test Results (Exploitation Phase)

The distribution of the residuals for the outcome feature botnets is seen to be centered at zero with narrow peaked curves and slightly skewed to the right. The distribution curve shows that the model performs well in predicting the database connections and the chat congestion features. The model

performs moderately in capturing the trends in the other features as the residuals are seen to be slightly skewed to the right.

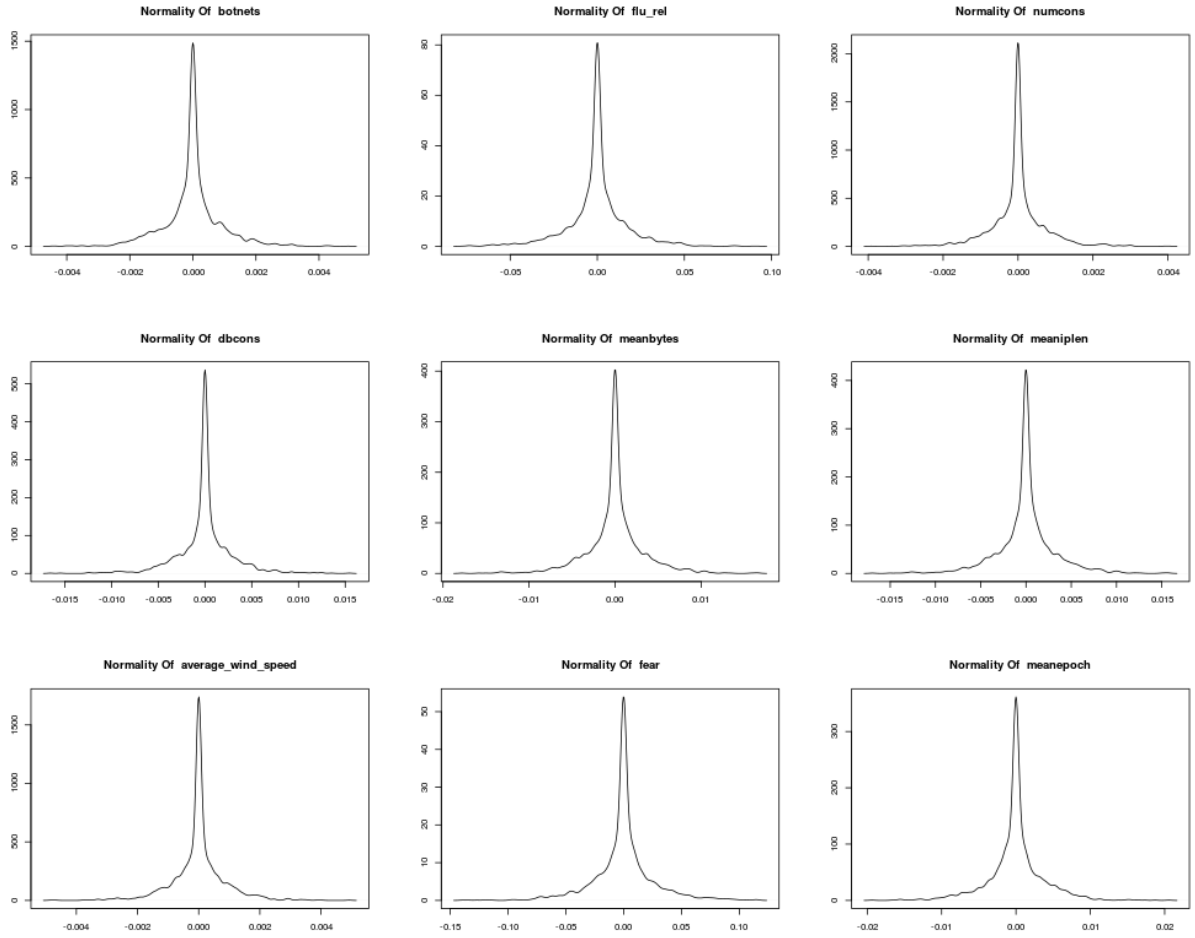

Figure 5-39: Density Distribution Of Residuals (Exploitation Phase)

Lastly, after observing the density distribution of the residuals, the researcher derives a normal probability plot for each residual from the structural model. To achieve this, the researcher calculates the cumulative probability of each residual using the formula:

$$P(i-th\ residual) = \frac{i}{(N+1)}$$

Where P is the cumulative probability of an observed residual, I is residual observation and N is the number of observations made within the given time period.

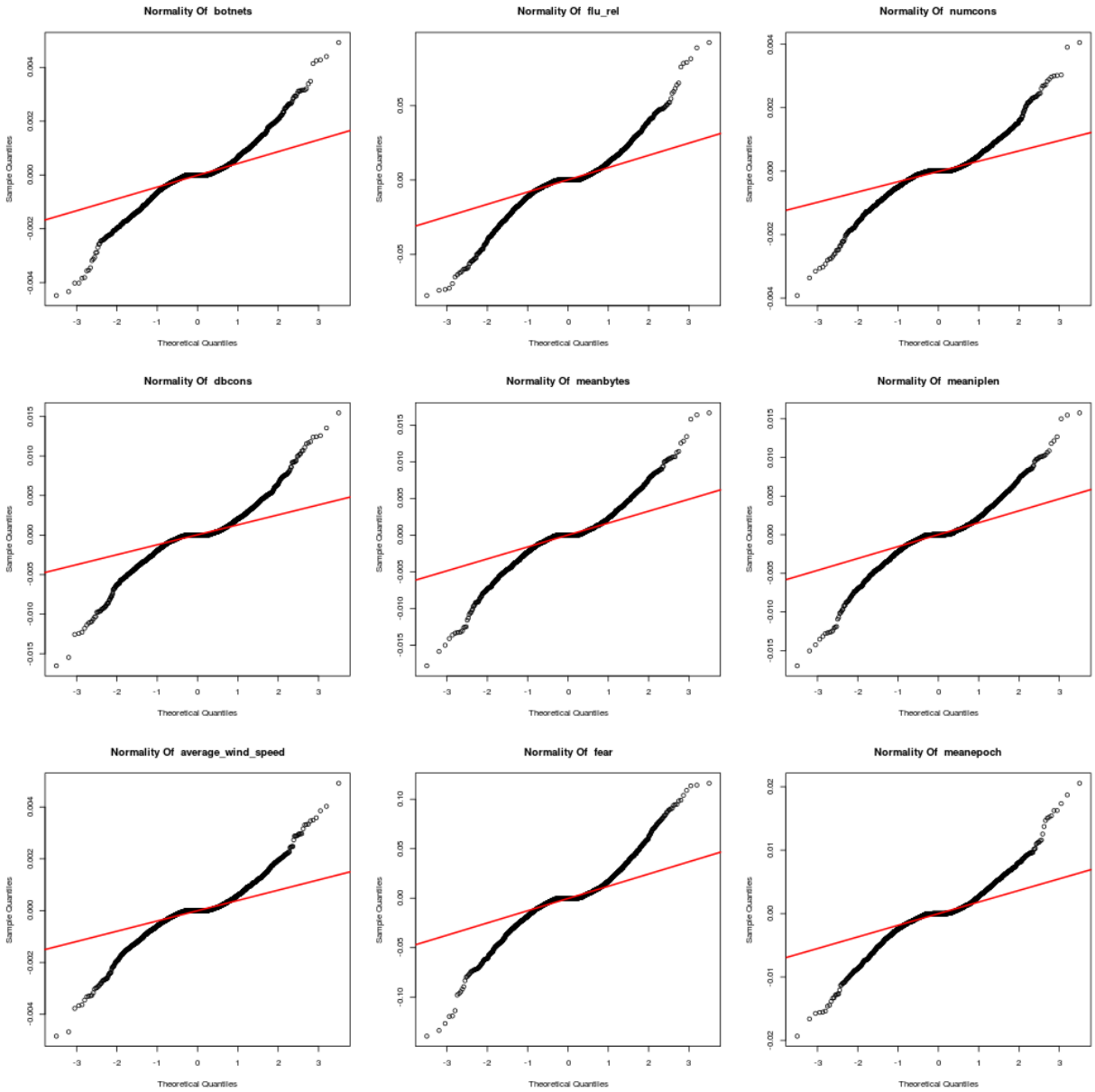

Figure 5-40: Normality Probability Plot of Residuals (Exploitation Phase)

The normal probability plot identifies departures from normality, non-linearity and outliers in a given distribution. Deviations from the straight line indicate a deviation from normality. A homoscedastic distribution is observed across all plots with reducing errors around zero. The S-Shape formed by data points in the plots indicate a shift towards normality. However, a few outliers are observed in all 9 features.

The performance of the VAR model is measured using structural multivariate time series model validation techniques outlined in the analytical frameowk. The model is validation using both in sample and out of sample prediction values. The in-sample residuals are derived by obtaining the differences between the actual observations and the predictions returned by the model. The results for the in-sample model performance is shown in the table below.

	MAE	RMSE	MPE	MAPE	MAD	FEATURE ACCURACY

Botnets	0.00	0.00	0.02	0.02	0.00	99.96
Flu Relatedness	0.01	0.02	1.33	1.33	0.00	98.92
Total Number Of Transmitted Packets	0.00	0.00	0.14	0.14	0.00	99.92
Database Connections	0.00	0.00	0.43	0.43	0.00	99.74
Mean Transmitted Bytes	0.00	0.00	0.05	0.05	0.00	99.87
Mean Ip Length	0.00	0.00	0.09	0.09	0.00	99.86
Average Wind Speed	0.00	0.00	0.00	0.00	0.00	99.97
Fear	0.02	0.03	-18.11	18.11	0.00	99.17
Mean Epoch	0.00	0.00	0.06	0.06	0.00	99.86
AVERAGE IN SAMPLE ERROR	0.00	0.01	-1.78	1.66	0.00	

Table 5-32: In Sample Model Performance Results (Exploitation Phase)

The model works well in predicting most of the features in the model. The model also performed accurately for features with smaller values of RMSE, MAD and MAPE. These results show that the model works well with expected data and fits the observations of interest.

Additionally, out of sample validations are also done for each feature Out of sample forecasts are obtained by forecasting the N-ahead data values where N=120 minutes i.e the last hour initially subtracted from the original dataset. The out of sample residuals are then derived by obtaining the differences between forecasts on the test datasets using the VAR model and the test data with 120 observations. The measures of out of sample performance values are presented in the table below.

	MAE	**RMSE**	**MPE**	**MAPE**	**MAD**	**FEATURE ACCURACY**
Botnets	0.50	0.00	0.02	1.78	0.06	88.44
Flu Relatedness	0.01	0.02	2.12	3.24	0.12	82.928
Total Number Of Transmitted Packets	0.03	0.00	0.14	-0.78	0.04	83.328
Database Connections	0.00	0.00	0.89	-4.02	0.01	90.426
Mean Transmitted Bytes	0.06	0.02	0.05	0.17	0.11	96.156
Mean Ip Length	0.04	0.01	1.01	0.13	0.03	83.772
Average Wind Speed	0.01	0.07	0.00	-0.03	0.02	73.606
Fear	0.02	0.03	12.11	13.98	0.07	72.982
Mean Epoch	0.19	0.30	0.06	1.14	0.02	89.56

AVERAGE TOTAL ERROR	0.10	0.05	1.82	1.73	0.10	

Table 5-33: Out of Sample Model Performance Results (Exploitation Phase)

The out of sample performance shows that the model works well in predicting new values of the endogenous feature. However, some features such as out of sample predictions for population fear of microblogging feeds are seen to exhibit a higher level of out of sample percentage error. Although these features are selected as being important for predicting the outcome variable, the features included in the model are not entirely efficient in predicting them. By averaging the performance measure across the various performance measurement techniques, the model records a total of 95% in sample accuracy and 90% out of sample accuracy for the outcome feature.

Following the performance test, the granger causality tests each feature combination with the optimum lag selected in the model order selection stage. The structural model can also be used to make inference about the direction or directions of causality between every pair of feature in the feature set. The granger causality test is set up with the null hypothesis of no causality between the feature 'x' (on the left-hand of the arrow) and the feature 'y' (on the right-hand of the arrow). The table below shows only causal links with significant P-values to reject the null hypothesis.

	Causal Relation	F-Statistic	P-Value
1.	Botnets --> Total Number Of Transmitted Packets	98.09035	1.08E-22
2.	Botnets --> Database Connections	50.94001	1.26E-12
3.	Botnets --> Mean Transmitted Packet Bytes	8.631804	0.003335
4.	Botnets --> Mean IP Length	9.264652	0.002362
5.	Flu Relatedness --> Database Connections	5.510133	0.018987
6.	Flu Relatedness --> Mean Transmitted Packet Bytes	60.27341	1.21E-14
7.	Flu Relatedness --> Mean IP Length	69.14968	1.51E-16
8.	Flu Relatedness --> Average Wind Speed	164.6685	1.66E-36
9.	Flu Relatedness --> Fear	32.46769	1.36E-08
10.	Flu Relatedness --> Mean Epoch Time	5.401285	0.020205
11.	Total Number Of Transmitted Packets --> Botnets	28.96232	8.10E-08
12.	Total Number Of Transmitted Packets --> Mean Epoch Time	5.631192	0.017722
13.	Database Connections --> Botnets	25.68134	4.33E-07
14.	Mean Transmitted Packet Bytes --> Botnets	16.26122	5.69E-05
15.	Mean Transmitted Packet Bytes --> Database Connections	7.856095	0.005106

16.	Mean Transmitted Packet Bytes --> Mean IP Length	16.39418	5.31E-05
17.	Mean Transmitted Packet Bytes --> Average Wind Speed	56.6992	7.14E-14
18.	Mean Transmitted Packet Bytes --> Fear	10.77685	0.001043
19.	Mean Transmitted Packet Bytes --> Mean Epoch Time	194.8797	1.22E-42
20.	Mean IP Length --> Botnets	14.09603	0.000178
21.	Mean IP Length --> Database Connections	7.847556	0.00513
22.	Mean IP Length --> Mean Transmitted Packet Bytes	15.0411	0.000108
23.	Mean IP Length --> Average Wind Speed	58.50122	2.92E-14
24.	Mean IP Length --> Fear	10.84633	0.001004
25.	Mean IP Length --> Mean Epoch Time	186.4537	6.14E-41
26.	Fear --> Mean Transmitted Packet Bytes	116.1097	1.79E-26
27.	Fear --> Mean IP Length	123.9961	4.07E-28
28.	Fear --> Average Wind Speed	144.5079	2.29E-32
29.	Fear --> Mean Epoch Time	12.00307	0.00054
30.	Mean Epoch Time --> Average Wind Speed	7.742741	0.005435

Table 5-34: F-Test Results for Granger Analysis (Exploitation Phase)

The granger analysis shows a total of 30 likely causal relations of 81 possible relations between the 9 features in the feature set. The outcome feature 'botnet' is seen to have plausible causal relations with 5 features on the physical dimension. A bi-directional causal relationship is also observed between botnet activities and mean frame length of transmitted bytes. Intra-dimensionally, there exists relations between features on the social dimension such as the population fear and the mean bytes transmitted on victim's network. Inter-dimensionally, there are observed causal relations between features within the physical dimension.

5.4.6 Stage 6: The Attack Phase

Stage 6 of the experiment integrates evidences leading from the physical dimension with evidences on the same dimension on a pre-defined time scale to spot indicators of a plausible detection of denial of service attempts on victim's network on the physical dimension of cyberspace. The rationale for conducting this stage of the experiment is to establish entanglements between events on the dimensions of cyberspace studied in this thesis using co-integral links between them at the 6th stage of the traditional kill chain model. The researcher uses series from the network layer of the physical dimension while propagating prior stages of the kill-chain already identified in previous stages and the social dimension in constructing a predictive model.

The endogenous variable selected for this phase characterizes a denial of service attack on victim's network as a ratio of the number of DNS requests on port 53 to the number of UDP packets sent and the total number of incoming connections on the victims network (Sans Institute, 2012). The 'dos' variable is therefore derived as a function of the total number of requests, the total amount of TCP packets and total number of SYN flags raised sent within selected time intervals on the victim's network (Sans Institute, 2012). It is estimated by scaling the total number of TCP Packets and the total number of SYN flags by the number of connections made and packets sent over the network.

$$\left(\frac{TCP\ Packets - SYN\ Flags}{Total\ Number\ of\ Connections}\right)$$

Equation 5-18: Characterizing DOS Attacks in Network Traffic

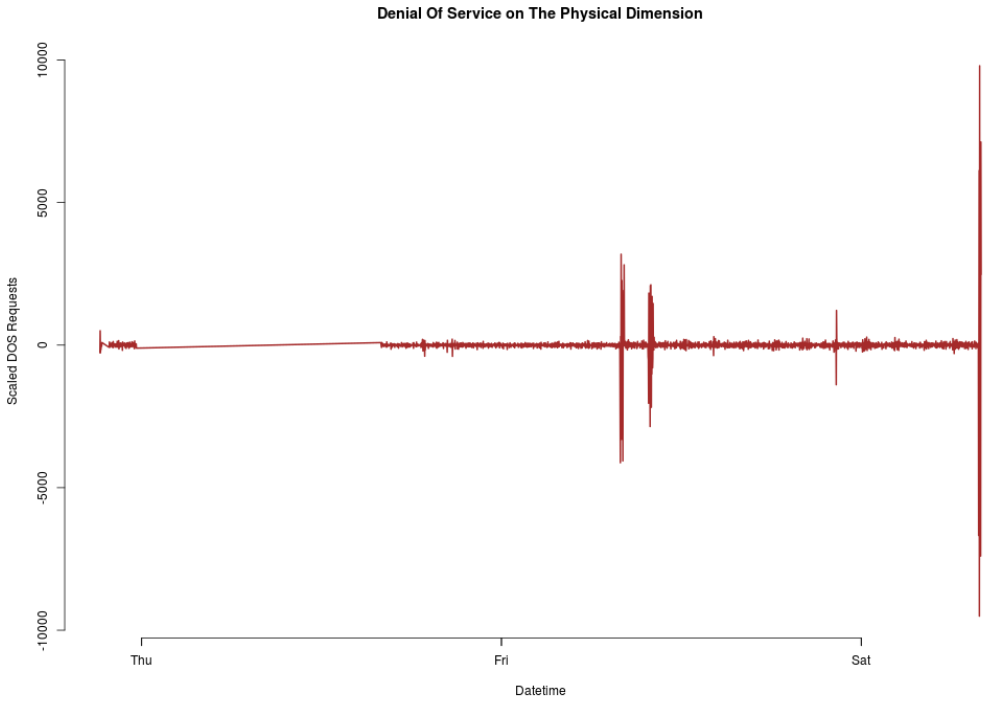

Figure 5-41: Denial Of Service on The Physical Dimension

The endogenous variable selected at this stage is the 'dos' which a denial of service attempt by external adversaries. As with the previous stage of the experiment, this feature is also arbitrary to represent multiple stages of the traditional kill-chain. Datasets from both dimensions were integrated on a similar timeline with the aim of predicting events of the last hour on the physical dimension. The merging of both datasets produced a single dataset with 74 features and 2524 observations. For testing and evaluation of developed model, the data was chronologically split into two, removing the last 120 observations (the two hours) from each feature in the dataset, as a separate test dataset. The out of sample reliability of the constructed model will be tested on observations from the last hour using an N-ahead prediction strategy where N=120. This is done to ensure reliability of out of sample prediction performance testing.

A simple intra-dimensional correlation analysis of the 74 features on both dimensions as shown in figure 5-4 below, reveals a slight correlation between features on both dimensions. For example, the

level of chat congestion and chatroom traffic is seen to have a slight correlation with the number of distinct connections to the network. Additionally, on the social dimension, emotion and opinion features are observed to be highly correlated with each other. The black boxes in the figure below groups features into hierarchical clusters. The recon, connection fail ratio, scanner, weaponization and injection features are propagated from the reconnaissance, weaponization, delivery and the C&C phases of the kill-chain. Sets of network and social features are observed to be highly correlated with each other such as ip addresses, ports, number of connections denied by intrusion detection system, total alerts raised by intrusion detection system, weaponization, chat room congestion. These correlation relationships were also observed in the previous stages of the experiment.

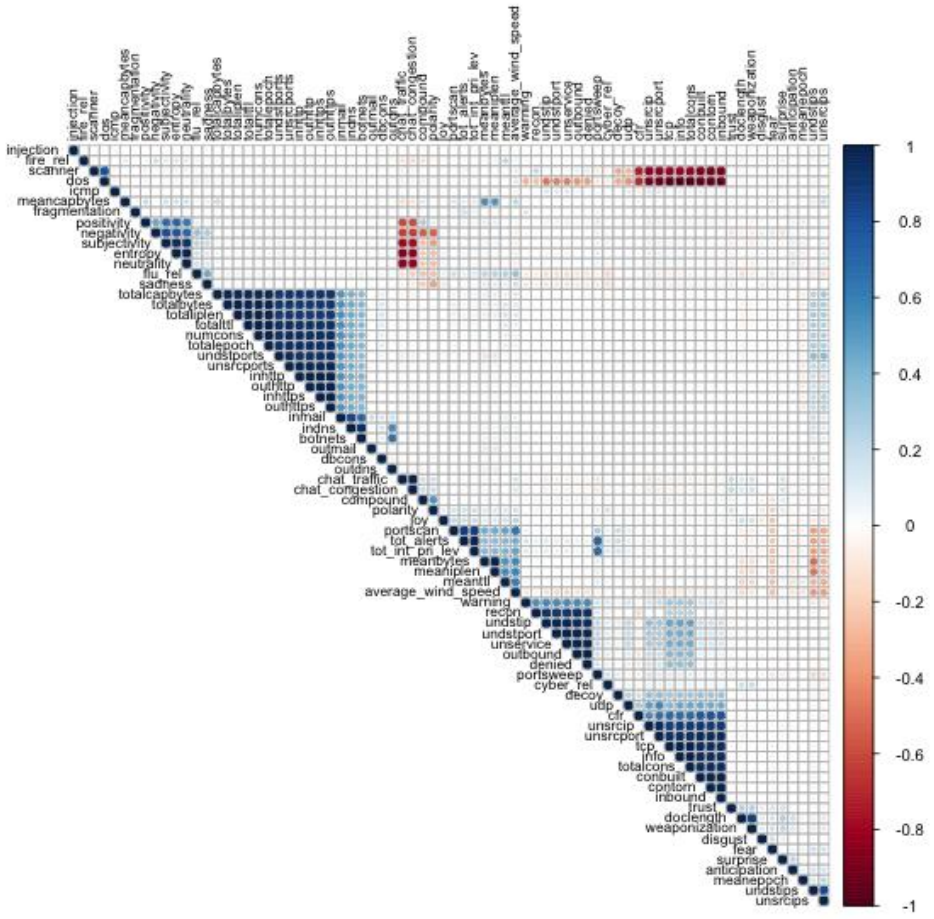

Figure 5-42: Intra-Dimensional Correlation Coefficients (Attack Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment relevant to the 6th phase of the experiment. The correlation coefficients are on a scale of +1 to -1 as shown by the color bar. Highly positively correlated pairs of variables tend towards +1 while highly negatively correlated pairs of variables tend towards -1. Multiple pairs of variables are seen to be highly positively or negatively correlated. For example, the group of features, Cyber relatedness, Entropy, chat size, chat room congestion and chat room traffic posted per minute are all seen to be highly positively correlated. Additionally, the group of features such as packet data features, reconnaissance features and intrusion detection features are also seen to be highly correlated with each other. Consequently,

one or more of pairs of features that are seen to be highly correlated with each other (where Pearson's correlation coefficient ≥ +0.65 or ≤ -0.65) are removed from the dataset. The correlation elimination process reduces the feature set by removing one or more of a set of highly correlated variables (Hall and Smith, 1998). The assumption is that information provided by highly correlated variables can be provided to the model by a single one of those variables therefore reducing redundancy to achieve a parsimonious model. The results of the filtered feature set after a correlation elimination produces a smaller feature set of 26 features as shown in the figure below, eliminating 48 highly correlated features.

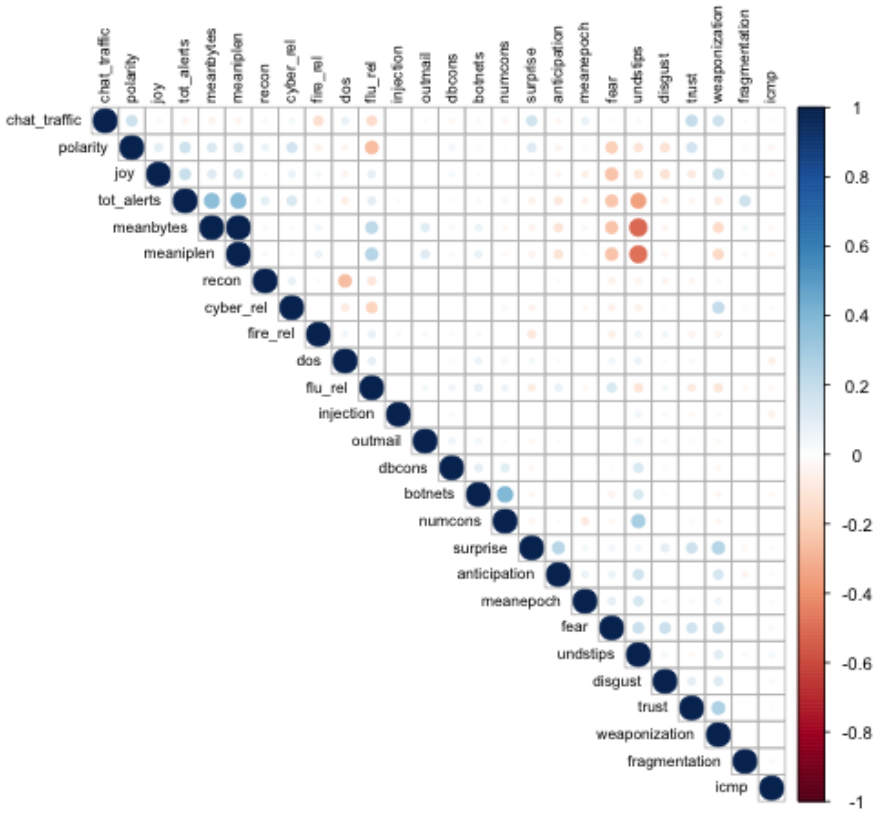

Figure 5-43: Correlation Coefficients Reduced Feature Set (Attack Phase)

The figure above shows the intra-dimensional correlation between features across the social and physical dimensions of the information space represented in this experiment relevant to the 6th phase of the experiment for the reduced feature set. This reduced set of 26 features including the endogenous variable, 'dos' are selected for further feature selection methods. To further reduce the number of features, the researcher applies a recursive features selection method to select only features that significantly reduce the error of prediction in the endogenous feature. The recursive feature selection stage further reduces the uncorrelated feature set to only those variables that minimizes the prediction error for the chosen endogenous variable 'dos'. To begin the recursive feature selection from the uncorrelated feature set, the researcher starts by building an ordinary least squares regression linear model with the 'dos' as the endogenous feature predicted by all other features in the reduced feature set above as shown in the figure above. The recursive feature selection searches for a model that optimizes the Akaike Information (AIC) and reduces errors. The recursive feature selection step further reduces the number of features leaving 10 features. The resulting features are shown in the table below and used in further analysis. Lastly, the researcher creates generic methods for calculating

the importance of each feature in predicting the outcome feature. This is used to examine the contribution of each feature in the dataset in predicting the outcome feature 'dos'. This last step in the feature selection phase is elimination by variable importance (Mehmood *et al.*, 2012). The absolute value of the t-statistic for each feature is used to compute the unique contribution of each input-feature to the model and a cut off of 1 is chosen (Noppamas, Seree and Kidakan, 2014). From the figure below, it is observed that there are 9 important features for predicting an attempted denial of service attack on victim's network. These features include botnet activities, chat traffic, reconnaissance, number of destination IP addresses active within a given time frame, the population fear, joy, flu relatedness, fire relatedness and cyber relatedness derived from microblogging feeds.

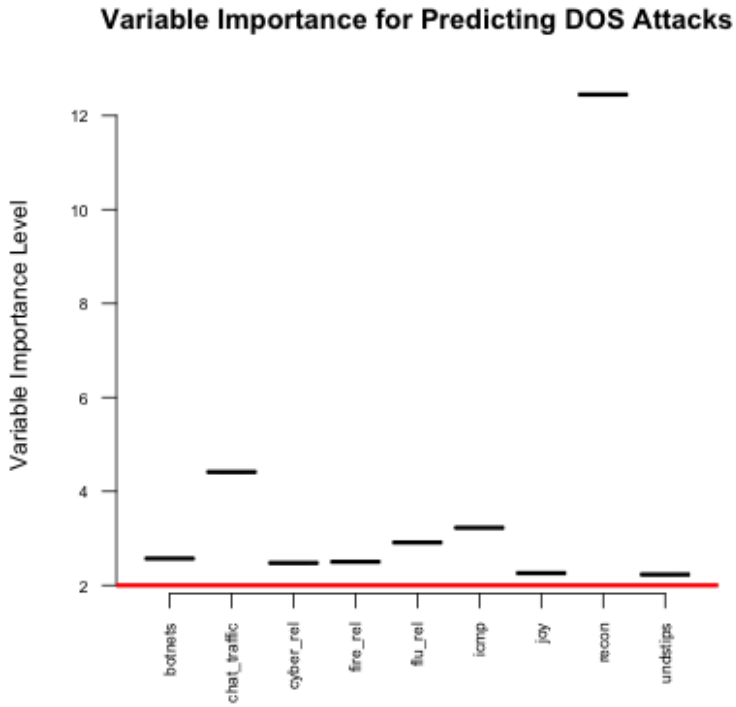

Figure 5-44: Variable Importance of Features (Attack Phase)

The figure above shows the relative variable importance for each feature in the final feature set. The last phase of the feature selection left the same 9 homogenous features as extremely useful for predicting the outcome feature. The recon variable is seen as the most useful feature for predicting the attack phase of the kill-chain (in a Denial of Service attack). Interestingly, event-related features from the social dimension such as the cyber-related of text discussions ('cyber_rel'), discussions about the fire event and flu spread ('fire_rel', and 'flu_rel') are also seen as significant indicators of this phase. The final features for the model are shown in the table 5-34 below.

The next phase of the analytical framework involves selecting appropriate orders for the intended model. This involves selecting an autoregressive order, testing for seasonality in each feature and selecting a moving average as discussed in the analytical framework.

The VAR model order selection phase seeks to select the appropriate significant lag at which features are serially correlated. The researcher uses an iterative solution with the maximum of 300 lags. At each iteration, the researcher builds a model for each equation in the VAR and measures the AIC. Finally, the researcher selects 216 lags (approximately 2 hours) as this is the lag length with the optimal AIC.

In addition to selecting appropriate VAR lag order, the researcher also tests each time series in the feature set for seasonality. In order to achieve this, the researcher employs a time series decomposition technique. As discussed in section literature review, time series decomposition works by splitting a time series into its three main components: the trend, the seasonal movement and the error terms. Seasonal patterns are expected to repeat with a fixed period of time. An additive decomposition method was chosen as the magnitude of the series does not increase with the series. After applying the Fourier transform for detecting seasonality, the botnets and database connections features shows an approximately 2 hours seasonal variation.

A stationarity and normality test are therefore conducted on the 10 features and results are tabled below. The results of the stationarity test show a significant p-value for rejecting the null hypothesis of non-stationarity in the augmented dickey fuller stationarity test. All selected features are seen to have a constant mean and variance over the selected period of time.

	Stationarity Test			Normality Test			
	P-Value	**Augmented Dickey Fuller Statistic**	**Stationary**	**Skewness**	**Kurtosis**	**Kolmogorov–Smirnov Test Statistic**	**K-S P-Value**
Dos	0.01	-11.56	True	-11.89	160.09	0.99	0.00
Recon	0.01	-9.03	True	15.75	269.98	0.52	0.00
Chatroom Traffic	0.01	-5.54	True	1.42	5.05	1.00	0.00
Botnets	0.01	-11.60	True	18.35	382.97	0.50	0.00
Cyber Relatedness	0.01	-8.98	True	0.58	1.05	0.50	0.00
Fire Relatedness	0.01	-7.23	True	1.81	5.63	0.50	0.00
Icmp Packets	0.01	-18.34	True	0.064	0.08	1.00	0.00
Flu Relatedness	0.01	-8.26	True	0.40	0.48	0.53	0.00
Joy	0.01	-9.34	True	0.27	0.43	0.50	0.00
Number Of Destination Ips	0.14	-3.03	False	0.64	-0.30	1.00	0.00

Table 5-35: Features' Stationarity and Normality Test Results (Attack Phase)

Additionally, the skewness and K-S tests show that the botnets, recon and dos features are slightly skewed and exhibiting varying peakedness. However, the kurtosis tests for the flu relatedness, cyber relatedness, fire relatedness, number of destinations ips and the chatroom traffic are seen to be fairly normal. Given that all features except the number of destination IPs were stationary, the Johansen's test for co-integration was not applicable at this stage of the experiment. However, for further investigation, the Engle and Granger test for co-integration was used to determine co-integration between each pair of features in the feature set. To achieve this, the researcher constructs a linear regression model between each pair of features and conducts a stationarity test on the residuals of the resulting model. The Engle Granger test shows no significant p-values for features in the system at the selected time period.

Given that all features are observed to be stationary and the absence of any co-integrating relationships at this stage of the experiment, a VAR model was fitted to the levels. The VAR with

order of 216 lags was estimated utilizing the OLS per equation in the model with 10 parameter OLS models where each feature is predicted by its on lags and lags of the other 9 features in the model. The density distribution plots below show the residuals from the fitted VAR(216) model.

The error terms i.e. the residuals are expected to be independently distributed across each equation and serially uncorrelated. The residuals are experimental errors derived by finding the difference between the observed data points and the predicted data points. To ensure that these assumptions are met, the researcher tests the hypothesis of white noise residuals using a normality test on the residuals of the VAR model. The stationarity and normality test results for the residuals are shown below.

	Stationarity Test			Normality Test			
	P-Value	**Augmented Dickey Fuller Statistic**	**Stationary**	**Skewness**	**Kurtosis**	**Kolmogorov–Smirnov Test Statistic**	**K-S P-Value**
Dos	0.01	-10.71	1.00	0.26	1.39	0.45	0.00
Recon	0.01	-15.47	1.00	-0.06	1.28	0.46	0.00
Chatroom Traffic	0.01	-16.37	1.00	0.03	1.29	0.41	0.00
Botnets	0.01	-17.87	1.00	-0.22	1.72	0.44	0.00
Cyber Relatedness	0.01	-13.99	1.00	0.11	1.67	0.41	0.00
Fire Relatedness	0.01	-18.35	1.00	0.10	1.75	0.42	0.00
Icmp Packets	0.01	-18.26	1.00	-0.02	1.31	0.43	0.00
Flu Relatedness	0.01	-13.08	1.00	-0.09	1.33	0.43	0.00
Joy	0.01	-17.39	1.00	0.06	1.37	0.39	0.00
Number Of Destination Ips	0.01	-13.77	1.00	-0.08	1.35	0.48	0.00

Table 5-36: Residuals' Stationarity and Normality Test Results (Attack Phase)

The distribution of the residuals for the outcome feature dos is seen to be centered at zero with narrow peaked curves and slightly skewed to the right. The distribution curve shows that the model performs well in predicting and in capturing the trends in most features as the residuals are seen to be slightly fairly normal and stationary.

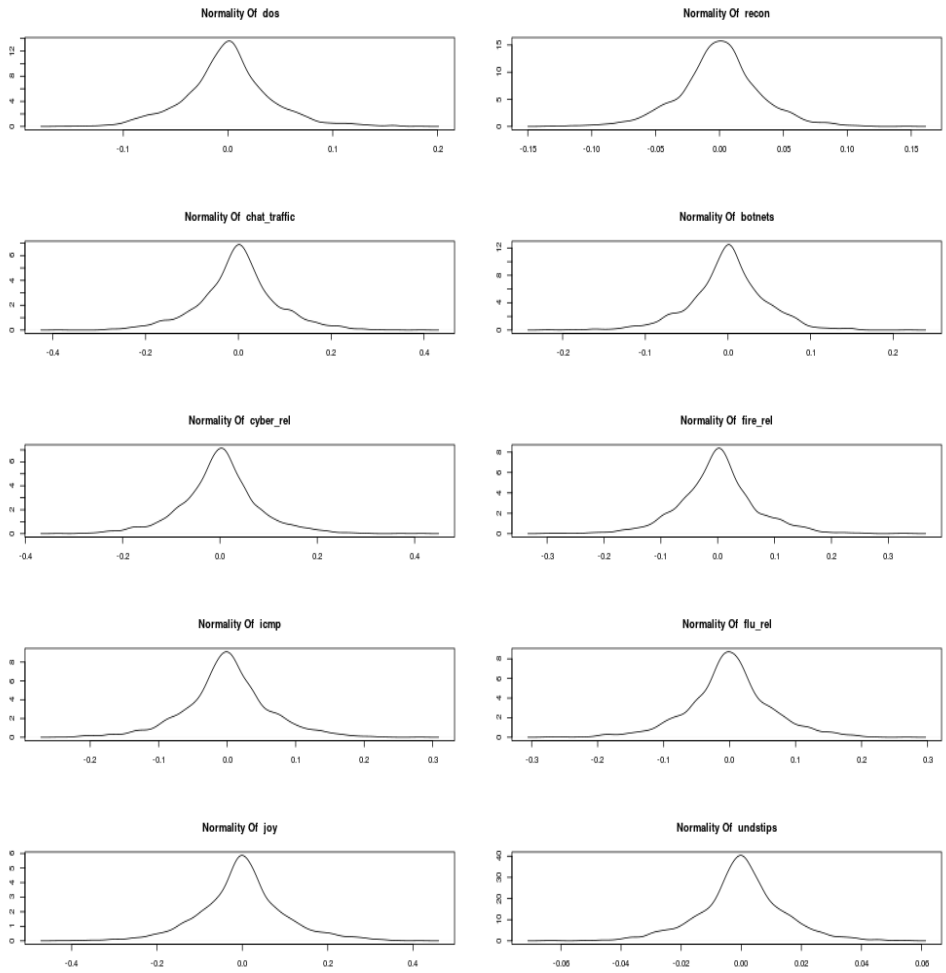

Figure 5-45: Density Distribution Of Residuals (Attack pHASE)

Lastly, after observing the density distribution of the residuals, the researcher derives a normal probability plot for each residual from the structural model. To achieve this, the researcher calculates the cumulative probability of each residual using the formula:

$$P(i-th\ residual) = \frac{i}{(N+1)}$$

Where P is the cumulative probability of an observed residual, I is residual observation and N is the number of observations made within the given time period.

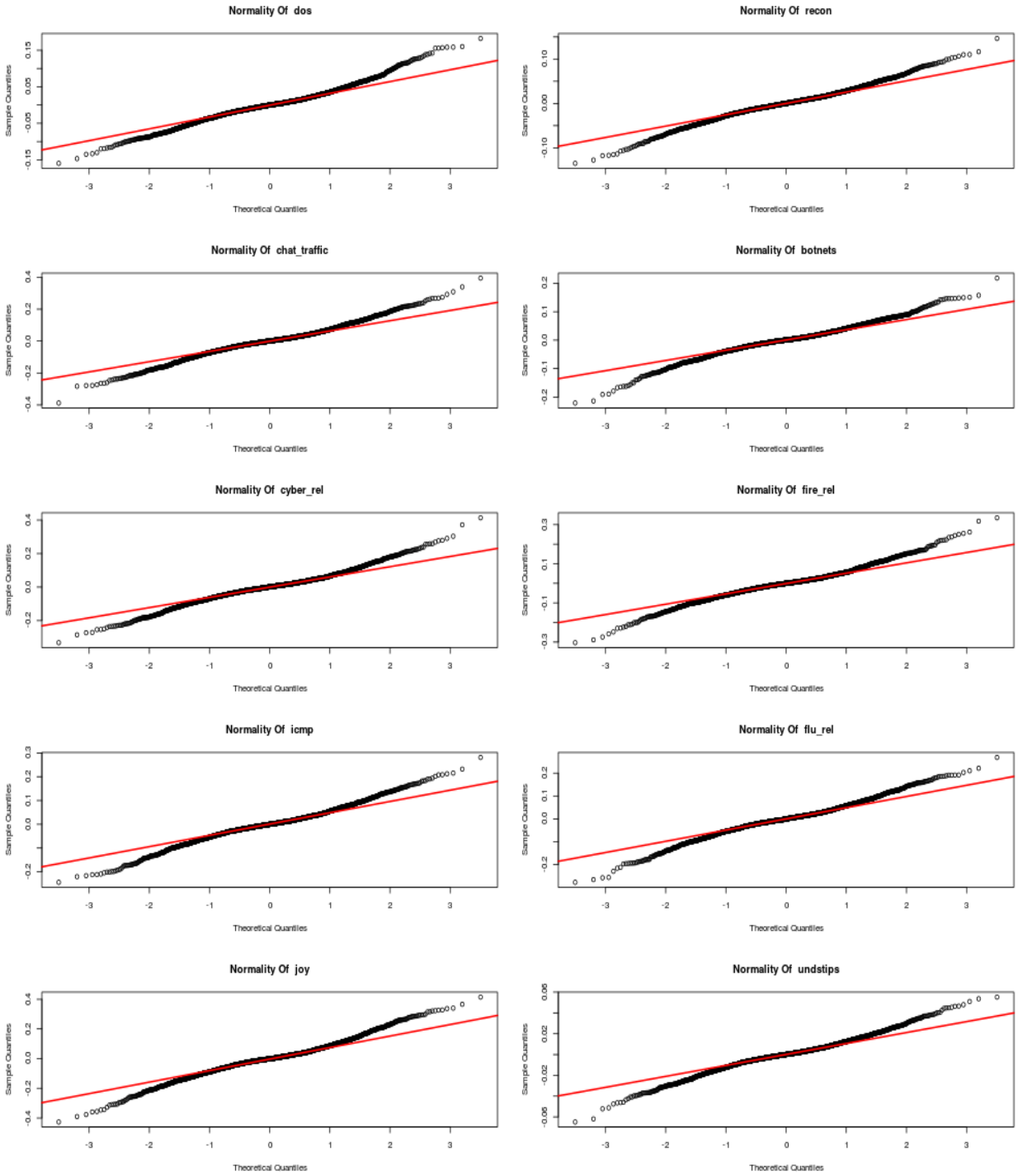

Figure 5-46: Normality Probability Plot of Residuals (Attack Phase)

The normal probability plot identifies departures from normality, non-linearity and outliers in a given distribution. Deviations from the straight line indicate a deviation from normality. A homoscedastic distribution is observed across all plots with reducing errors around zero. The S-Shape formed by data

points in the plots indicate a shift towards normality. However, a few outliers are observed in all 10 features.

The performance of the VAR model is measured using structural multivariate time series model validation techniques outlined in the analytical framework. The model is validation using both in sample and out of sample prediction values. The in-sample residuals are derived by obtaining the differences between the actual observations and the predictions returned by the model. The results for the in-sample model performance is shown in the table below.

	MAE	RMSE	MPE	MAPE	MAD	FEATURE ACCURACY
Dos	0.03	0.04	1.14	1.14	0.00	98.11
Recon	0.02	0.03	-25.07	25.07	0.00	98.88
Chatroom Traffic	0.06	0.08	91.26	91.26	0.00	60.58
Botnets	0.03	0.05	40.95	40.95	0.00	82.03
Cyber Relatedness	0.06	0.08	-5.72	5.72	0.00	97.25
Fire Relatedness	0.05	0.07	78.18	78.18	0.00	66.36
Icmp Packets	0.05	0.06	-17.90	17.90	0.00	97.84
Flu Relatedness	0.05	0.06	-6.93	6.93	0.00	97.78
Joy	0.07	0.10	9.33	9.33	0.00	92.78
Number Of Destination Ips	0.01	0.01	0.49	0.49	0.00	99.32
Total Average Error	0.04	0.06	16.57	27.70	0.00	

Table 5-37: In Sample Model Performance Results (Attack Phase)

The model works well in predicting most of the features in the model. The model also performed accurately for features with smaller values of RMSE, MAD and MAPE. These results show that the model works well with expected data and fits the observations of interest.

Additionally, out of sample validations are also done for each feature Out of sample forecasts are obtained by forecasting the N-ahead data values where N=120 minutes i.e the last hour initially subtracted from the original dataset. The out of sample residuals are then derived by obtaining the differences between forecasts on the test datasets using the VAR model and the test data with 120 observations. The measures of out of sample performance values are presented in the table below.

	MAE	RMSE	MPE	MAPE	MAD	FEATURE ACCURACY
Dos	10.68	13.59	-93.88	93.88	0.09	51.29
Recon	1.96	2.50	2130.88	2130.88	0.02	5.83
Chatroom Traffic	2.24	2.90	102.48	102.48	0.02	75.58
Botnets	2.27	2.98	-83.88	24.74	0.02	70.65
Cyber Relatedness	2.27	3.04	566.58	566.58	0.02	66.68
Fire Relatedness	2.29	2.98	-22641.30	22641.30	0.02	55.41
Icmp Packets	2.25	2.95	-726.61	726.61	0.02	69.58
Flu Relatedness	2.42	3.21	-226.18	226.18	0.02	68.69
Joy	2.27	3.08	201.74	201.74	0.02	71.20
Number Of Destination Ips	2.15	2.92	-171.61	171.61	0.02	89.83

Total Average Error	3.08	4.01	-2094.18	2688.60	0.03	

Table 5-38: Out of Sample Model Performance Results (Attack Phase)

The out of sample performance shows that the model works well in predicting new values of the endogenous feature. However, some features such as out of sample predictions for population fear of microblogging feeds are seen to exhibit a higher level of out of sample percentage error. Although these features are selected as being important for predicting the outcome variable, the features included in the model are not entirely efficient in predicting them. By averaging the performance measure across the various performance measurement techniques, the model records a total of 95% in sample accuracy and 90% out of sample accuracy for the outcome feature.

Following the performance test, the granger causality test each feature combination with the optimum lag selected in the model order selection stage. The structural model can also be used to make inference about the direction or directions of causality between every pair of feature in the feature set. The granger causality test is set up with the null hypothesis of no causality between the feature 'x' (on the left-hand of the arrow) and the feature 'y' (on the right-hand of the arrow). The table below shows only causal links with significant P-values to reject the null hypothesis.

	Causal Relationship	F Statistic	P-Value
1.	**Denial Of Service Attempt --> Recon**	93.56279	9.76E-22
2.	**Denial Of Service Attempt --> Cyber Relatedness**	6.425258	0.011314
3.	**Recon --> Cyber Relatedness**	9.147256	0.002517
4.	**Recon --> Flu Relatedness**	5.330605	0.021039
5.	**Chatroom Traffic --> Joy**	5.057547	0.024609
6.	**Botnets --> Number Of Destination Ips**	7.495017	0.006233
7.	**Cyber Relatedness --> Denial Of Service Attempt**	17.69884	2.68E-05
8.	**Cyber Relatedness --> Recon**	17.27046	3.36E-05
9.	**Cyber Relatedness --> Flu Relatedness**	15.77563	7.34E-05
10.	**Cyber Relatedness --> Number Of Destination Ips**	8.737648	0.003147
11.	**Fire Relatedness --> Chatroom Traffic**	33.67483	7.37E-09
12.	**Fire Relatedness --> Flu Relatedness**	28.6925	9.29E-08
13.	**Fire Relatedness --> Joy**	13.69332	0.00022
14.	**Flu Relatedness --> Denial Of Service Attempt**	13.5687	0.000235
15.	**Flu Relatedness --> Recon**	6.928613	0.008537
16.	**Flu Relatedness --> Fire Relatedness**	6.064838	0.01386
17.	**Flu Relatedness --> Joy**	15.85585	7.04E-05

18.	**Flu Relatedness --> Number Of Destination Ips**	14.64731	0.000133
19.	**Joy --> Cyber Relatedness**	8.683932	0.003241
20.	**Joy --> Flu Relatedness**	6.379723	0.011607
21.	**Joy --> Number Of Destination Ips**	27.164	2.03E-07
22.	**Number Of Destination Ips --> Botnets**	63.50503	2.45E-15

Table 5-39: F-Test Results for Granger Analysis (Attack Phase)

The granger analysis shows a total of 22 likely causal relations of 90 possible relations between the 10 features in the feature set. The outcome feature 'dos' is seen to have plausible causal relations with 3 features, 1 on the physical dimension and 2 on the social dimension. A bi-directional causal relationship is also observed between denial of service attempts and the spread of epidemic flu in the population. Additionally, the 3 major events of the scenario design are observed to intra-dimensionally linked to the denial of service attempts on the network layer. Intra-dimensionally, there exists relations between features on the social dimension such as the population joy and the number of destination IPs communicating on victim's network. Inter-dimensionally, there are observed causal relations between features within the physical dimension such as the number of destinations ips communicating on victim's network and botnet activities. There is also a bi-directional relationship observed between these two features.

5.5 CONCLUSION

This chapter has presented the practical implementation of the entangled cyberspace framework with an alignment to the theory developed in chapter 2 of this report. It presents the experimental test environment as well as the proposed analytical and implementation strategy for the framework. It showed how data used in evaluating the theoretical framework was acquired, transformed and analysed to yield supporting results. The accuracy and errors for each model developed at each phase of the experiment was also measured and reported and their impact in the usefulness of the framework. The next chapter discusses the results of the analysis and therefore the findings of this research.

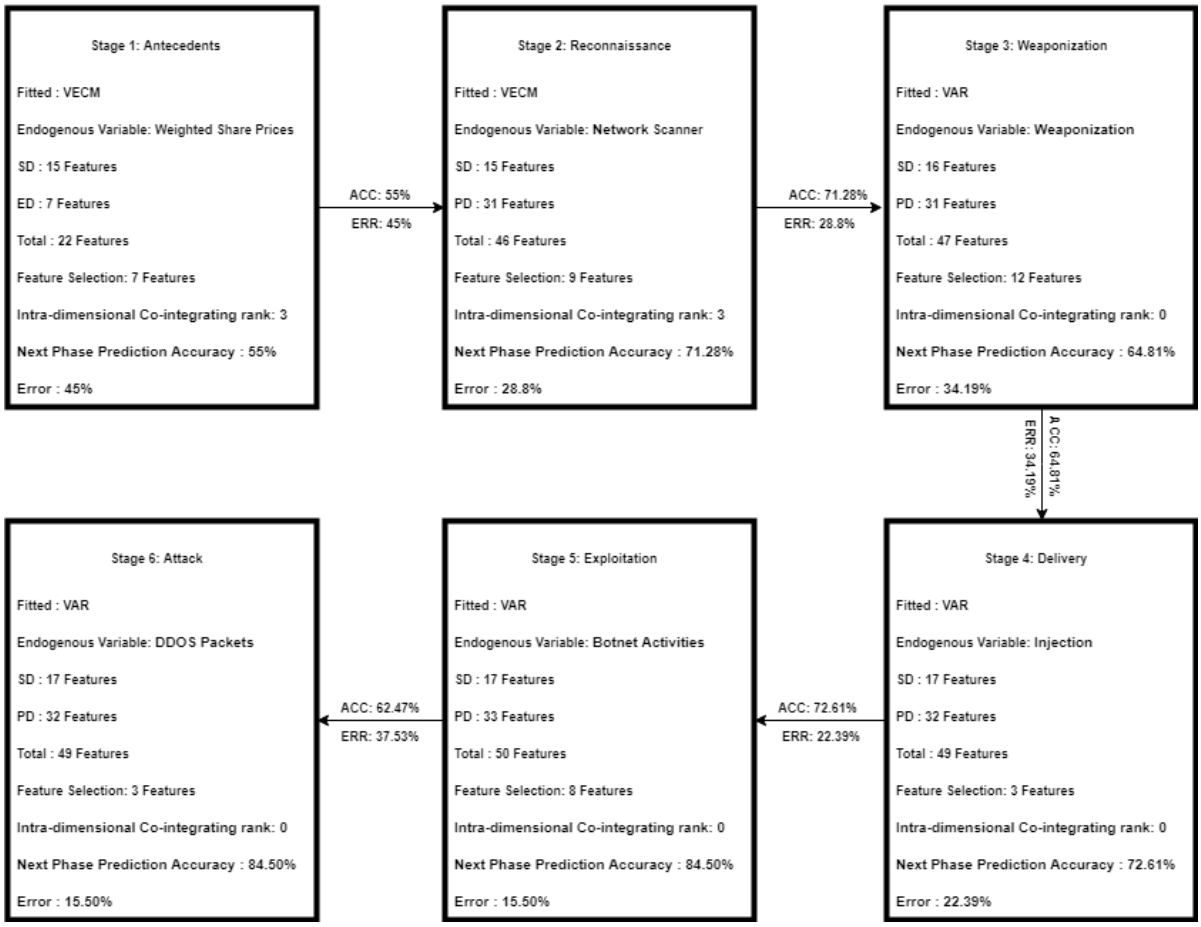

Figure 5-47: Summary Of Experiments

The figure above depicts the summary of experiments for each stage of the kill-chain. The approach presented in this research are experimentally tested in pre-empting each stage of the kill-chain. We performed an in sample and out-sample testing of how well the approach is able to pre-empt the next phase of the kill-chain using features identified across various dimensions of the information space. The ability of our approach to pre-empt the next phase of the kill-chain is documented as the difference of our predicted values from our actual values. We estimate the error recorded at each phase using metrics discussed in section 5.3.11.2.

The ability of the approach to record a reduction in the propagated error across the phases demonstrates that the approach is able to detect features that are useful for characterizing various phases of the kill-chain. However, the error recorded at the 'weaponization' phase is seen to have increased approximately 7% from the last phase. These may be due to the fact that the approach has limited access to representative data at this phase.

Furthermore, the determination of the level of 'acceptable error' is strictly determined by the domain of analysis and an acceptable threshold set by some criteria determined by the cyber analyst.

6 CHAPTER 6 DISCUSSION OF FINDINGS

6.1 INTRODUCTION

This empirical results and analysis chapter discusses critically, the results from the experimental design and testing of the Entangled Cyberspace Theory using the scenario events. The experiment was conducted in six stages representing the phases of the generic cyber-attack kill-chain presented by Lockheed Martin (Hutchins, 2011). This chapter is organised as follows:

- The events within the scenario are described in the next section.
- Results of event analysis from each evidence source included in the scenario are discussed in sections 5.3, 5.4 and 5.5.
- Results of the intra-dimensional co-integration analysis across the three dimensions represented are presented in section 5.6.
- The physical-social-economic kill chain is presented in section 5.7.
- Results of the intra-dimensional causal analysis are presented in section 5.8.
- Finally, the evaluation of the hypothesised theoretical framework is presented in section 5.9.

This chapter aims at addressing the alignment of the research experiment and results with the research aims and objectives outlined in Chapter 1 of this report. This section also validates the experimental design as sufficient for proving the hypothesised theory and answering the research questions of entanglements between dimensions of cyberspace. The results of the analysis are explored and reported systematically following the research design and approach discussed in Chapter 3 of this report. To achieve these stages of intra-dimensional analysis in cyberspace, the researcher independently analyses results from the cyber-attack kill-chain phases as they are perpetrated across the various dimensions of cyberspace.

6.2 THE SCENARIOS IN CONTEXT

The scenarios used in this research are a combination of events based on the IEEE VAST 2011 visual analytics challenge (Cook *et al.*, 2011) and real-world events in the financial sector. The individual events were designed to simulate a flu epidemic spread, a fire accident in the population of a fictitious metropolitan city, a denial of service attack on a fictitious company and a stock market crash in the share prices of the company. Each event included in the scenario is designed to characterise activities on a single dimension of cyberspace. The events in the microblogging feed, the cyber-attack and the stock market crash are designed to characterise activities on the social, physical and economic dimension respectively.

6.2.1 Activities on the Social Dimension

Health professionals have noticed a drastic increase in the number of flu-related illnesses. Observed symptoms are reported to include fever, chills, sweats, swollen lymph nodes, fatigue, aches, pains coughing, breathing difficulty, nausea and vomiting. Microblog feeds collected from various devices owned by the population within the metropolitan area of interest are used to characterise various events occurring in the population. A deep analysis of the microblogging feeds identifies three main events: an Epidemic Spread, a Truck Accident characterized by reports of a fire and discussions related to a cyber-event.

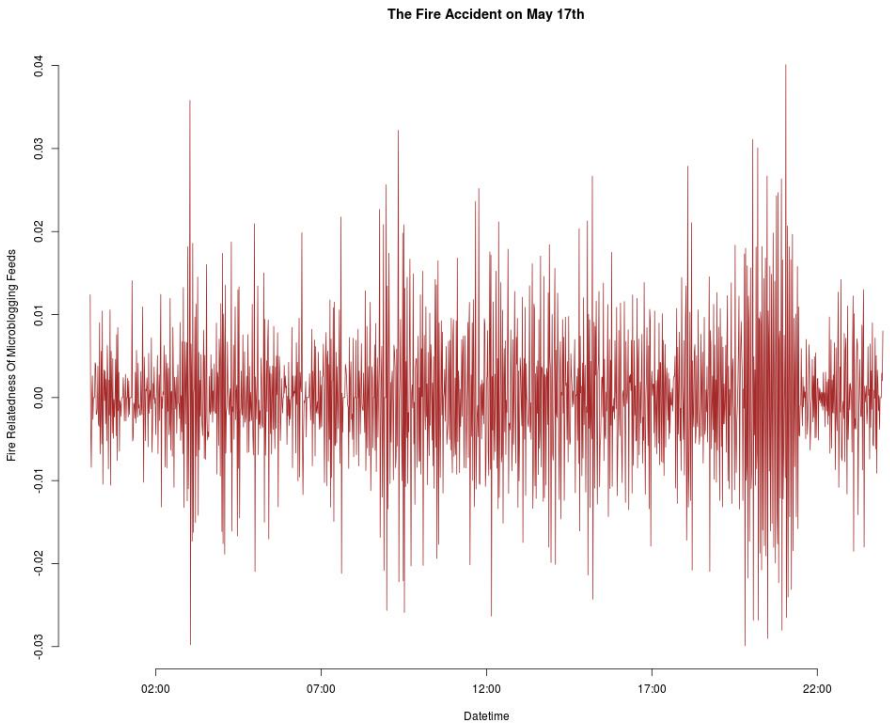

Figure 6-1: Scenario Event - Fire Accident on the 17th of May

The results of applying a fire-related lexicon to the microblogging feed provide a detailed quantification of feeds based on the truck accident event. On the 17th of May, there are reports of a truck accident within the microblogging feeds. The truck accident is reported in 500 microblog feeds on the 17th of May within the observed period. The incidents as depicted by the figure above shows major discussions at around 2 am, 8 am and 9 pm during the day.

Additionally, between 8 am on the 17th of May, and 8 am on the 20th of May increased reports of flu-related symptoms are observed within the microblog feeds. Similarly, the results of applying the flu-related wordlist to the microblogging feeds provide a detailed quantification of feeds based on the flu epidemic event. Infected individuals are seen to be randomly across the population area. From the figure below, the spike in reports on flu-related symptoms is observed by the drastic increase in flu relatedness of microblogging feeds beginning on the 17th of May up until about the 20th of May.

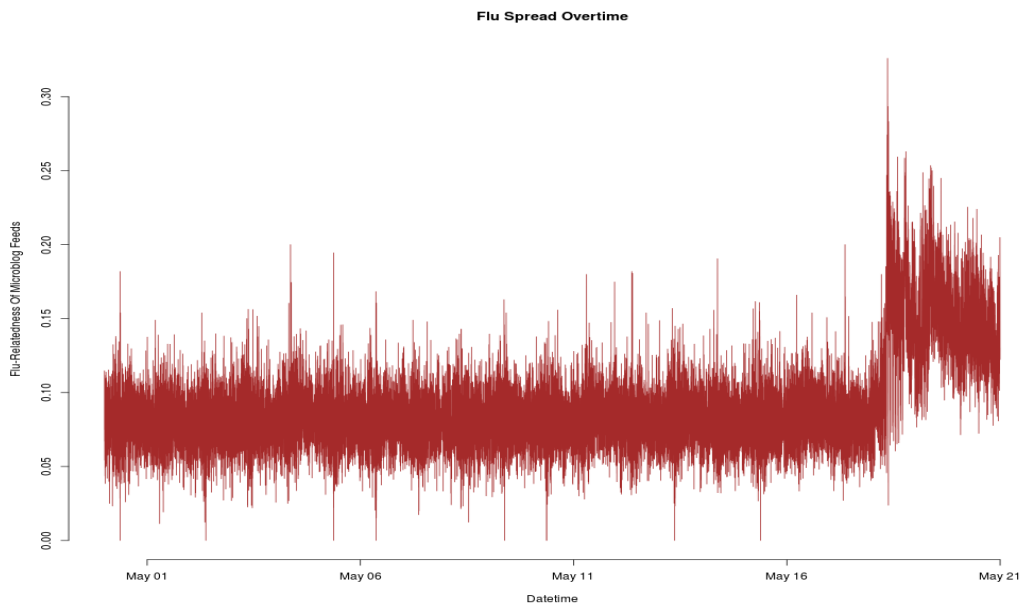

Figure 6-2: Scenario Event - Flu Spread Within Population

Furthermore, the spread of this epidemic or flu in the population is also seen to be related to the average wind speed of the area from where the feed was collected. The figure below shows data points for only days within the 16th and 20th of May. The average wind speed for these three days was either 5, 8 or 9. It is observed that on days with higher records of wind speed there is a higher drift in the observed levelled of flu relatedness in the microblogging feeds.

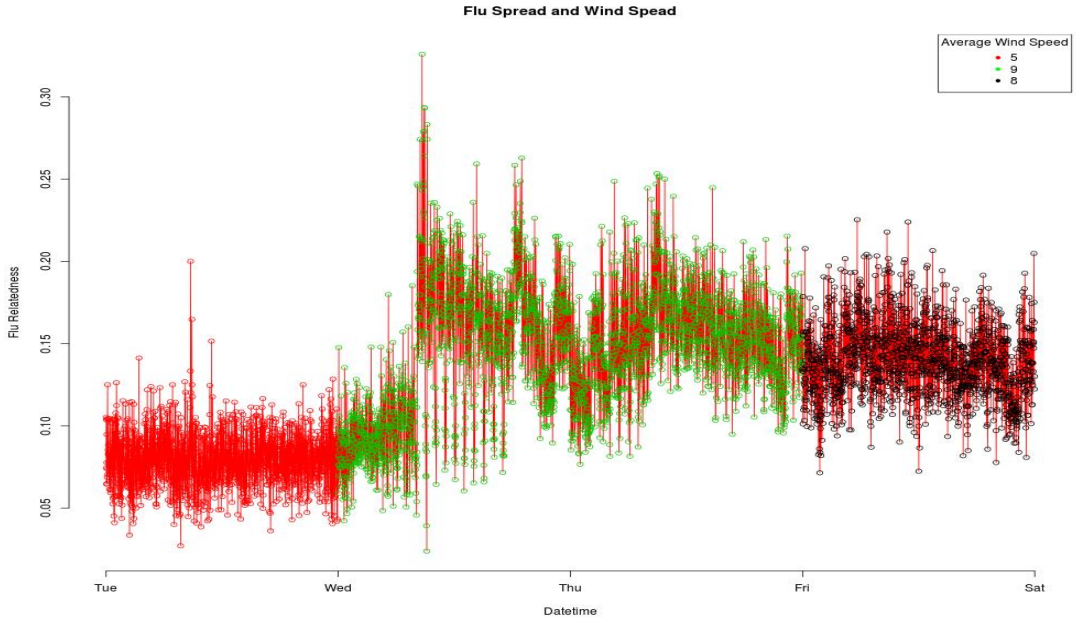

Figure 6-3: Scenario Event - Flu Spread VS Wind Speed Over Time

6.2.2 Activities on the Physical Dimension

The company in the scenario is a fictitious company based in the U.S. The company operates a corporate network with a cyber-team responsible for managing all aspects of the corporate network. The company is recently concerned with the level of cyber situational awareness of its employees and is dedicated to developing this capacity. To achieve this, the company analyses log files including

intrusion detection logs, firewall logs and network traffic logs to identify events of concern and develop a holistic cyber situational awareness strategy. The network architecture of the company's network is provided in chapter 4 of this report. The critical assets identified in the network are high priority nodes essential to the business processes of the company. A thorough analysis of the timeline of events on the physical dimension shows a streak of cyber incidents of interest beginning on the 11th of April 2011 and ends on the 15th of April 2011. The dates and times for the earliest detection of activities on each day is shown in the table below.

SN	Event	Description	Day	Evidence Source
1	Nessus Scan	The scan shows that the remote host runs a version of Windows known to be affected by multiple buffer overflow vulnerabilities when viewing WMF files. This may allow adversaries to execute arbitrary code on the remote host.	10:16 am 11th May 2011	Nessus Security Log
2	Attempted Denial of Service Attack	External systems are attempting to disrupt communications with the web server. This is an attack on the corporate web server to attempt to disrupt communications.	11:39 am 13th May 2011	Firewall Logs
3	Port Scan	All Freight computers begin port scanning other systems on their subnet (which is why this is not detected in the Firewall logs). This suggests a problem within the All Freight network, such as a worm.	11:15 am 13th May 2011	PCAP logs, IDS Logs
4	Port Scan	All Freight computers continue port scanning other systems on their subnet (which is why this is not detected in the Firewall logs).	09:01 am 14th May 2011	PCAP logs, IDS Logs
5	Port Scan	This port scan attack is from the unpatched workstations which are port scanning internal machines across subnets	10:56 am 14th May 2011	PCAP logs, IDS Logs
6	SMTP Email	An email exchange between two AFC employees. Represents the initiation of a socially engineered attack. Additional information about the activity can be found in the PCAP file during this time.	11:23 am 14th May 2011	PCAP Logs

7	Remote Desktop Connection	Remote Desktop Connection set to alert by emergency if detected from the outside. Remote Desktop connections from outside the network are prohibited by policy, but the network administrator has not blocked the connections on the firewall.	01:31 pm 14th May 2011	Firewall Logs
8	Authentication to Domain Controller	This is an authentication log for event ID 4634 to the web server corresponding to the unblocked external login	01:31 pm 14th May 2011	Security Logs
9	New Computer Powered on	New adversary infiltrated victim network.	01:23 am 15th May 2011	Firewall Logs, PCAP Logs

Table 6-1: Timeline of Events on the Physical Dimension

For this analysis, the activities shown in the table above are summarized as three main cyber-incidents on the victim's network: a) Denial of service attacks, b) Port Scan attacks and c) Social Engineering attack. Port scans are seen to happen from within the network which may suggest infected machines scanning for other hosts on the network to spread infection. Denial of service attacks are seen to be from external systems attempting to take down the company's corporate web server running on the host machine (see network architecture section). Finally, the SMTP, remote desktop control and authentication attacks sum up to a strategic social engineering attack designed to infiltrate the victim's network. The adversary uses this attack to disguise an externally controlled computer as a legitimate node on a victim's network.

6.2.3 Activities on the Economic Dimension

Events on the economic dimension lead up to a stock market crash in the stock prices of Delish Corporation. The timeline of share price fluctuations on the economic dimension takes place throughout one year. Significant price increases and price drops are observed at equal time interval deviating from the series equilibrium. The stock price crash happens before the activities on the physical and social dimensions and acts as antecedents to the events on these dimensions. The experiments show that the timeline of events on the economic dimension can be directly linked to events on the physical dimension, specifically on the network dimension at the first stage of the traditional kill-chain model. The predictive accuracy of Network scan and reconnaissance features on the network layer of the physical dimension using the weighted prices from the economic dimension is seen to be highly significant.

OBJECTIVE 1: To characterize cyberspace as a collection of multiple dimensions that are part of a whole

6.3 PHASED ANALYSIS

This section discusses the causal relationships observed between features across the various dimensions of cyberspace. Each stage of the experiment conducts a simple causal analysis with features across various dimensions. Although not all features with significant causal relationships were significant in predicting the outcome feature at that stage, the observed features with significant causal relationships represent an inter-connection of events at that stage of the kill chain. The

subsequent sub-sections discuss the results of these analysis at each stage of the kill-chain experimentation.

Stage one, the antecedent phase highlights causal relationships between user emotion features, event features (such as the epidemic spread within the population) and financial features (such as the weighted price index of company prices). Additionally, these features are also seen to exhibit a feedback system where user emotion features and event features are simultaneously predictive of each other. This phenomenon speaks to the interdependence of this human emotion on real-world events as identified in the experiment.

Stage two of the experiment analyses events on the network layer of the physical dimension to pre-empt an active reconnaissance. Vulnerability detection and intrusion detection features ae not only seen to be co-integrated but also causally linked within the given timeframe. Vulnerability detection features such as port scans and ping sweeps lead up to the detection of intrusions such as the attempted denial of service attack at 11.39am on the 13th of May. Additionally, the causal feedback loop between these features indicates a cyclic pattern between adversaries searching for exploitable vulnerabilities on victim's network and attempted attacks on the network such as the DOS and the outright infiltration of victim's network. Furthermore, there is also a minimal causal relation identified between the population's sadness and cyber-related gists on the social dimension. These features are seen to affect the population's perspective towards events occurring in the experiment.

Stage three of the experiment tracks the weaponisation phase of the Physical-Social-Economic kill-chain on the social dimension of cyberspace. This event is measured by lexical analysis of user texts based on a cyber weaponization wordlist. Consequently, we observe causal relationships between this feature and other features on the social dimension. For example, the emotion features and user opinion features such as polarity of user gists, are seen to be causal with each other. More importantly, the amount of information flowing through and from victim's network (as measured in the number of data packets through the network) is also seen to be linked to event features (Epidemic Flu Spread) and user emotion features (Anticipation, trust and fear).

Stage four of the experiment integrates evidence from the physical and social dimension to spot indicators of a cyber-weapon injection into the victim's network. This event is characterized as a function of the total number of packets sent over the network, the total number of incoming and outgoing database connections on the network and the total number of identified malicious HTTP requests. This stage of the experiment identifies only two significant causal relationships:

Causal relationships are observed between the social and physical dimensions and within the social dimension at the injection phase. The derived injection feature is seen to highly co-integrated and causal of the chatroom congestion on the social dimension. Inter-dimensionally, there is one observed significant causal relationship between the chatroom traffic and the population fear.

Stage five of the experiments tracks botnets in the operating network using evidence from the social and physical dimensions. The network layer characterizes this event as a function of time aggregated DNS requests and UDP connections in the operating network. Here, the research identifies approximately 30 significant causal relations between features on the physical and social dimensions with a 60% bi-directional causation. As expected, there are observed circular causal relationships between the derived botnet features and other features on the network layer like the total number of transmitted packets, the mean number of bytes transmitted within a given time frame and the total number of database connections on the network. Additionally, there are significant causal relations between features on the physical and social dimensions such as the epidemic spread and population fear as indicated on the social dimension and the number of database connections and the mean number of data bytes transmitted over the network within a given time frame.

6.4 PREDICTIVE FEATURES WITHIN DIMENSIONS

The causal analysis leads to the identification of useful predictive features on each dimension at each phase of the proposed kill-chain. These are features identified by this research as early indicators of kill-chain phases. The identification of these features across the dimensions solidifies the research theory of entanglements between dimensions of cyberspace, put in the context of a traditional kill-chain.

6.4.1 Predictive Features on the Social Dimension

Through the course of the experiment, several features from the social dimension have been observed to be significantly important to the entangled cyberspace eco-system. The scenario and experimental design use three sets of features to characterize events on the social dimension: Emotion-based features, Opinion–based features and Event-based features. This research uses a bag-of-words approach with various context wordlists to characterize various events from microblogging discussions. These characterizations track discussions on certain events such as the epidemic flu spread, a perceived creation of a cyber-weapon and a truck accident on the social dimension. These features are referred to in this research as 'event-based' features. Similarly, the emotions features include joy, fear, anticipation, surprise, sadness, trust while the emotion-based characterise the sentiments of discussions.

SN	Feature	Stage1	Stage2	Stage3	Stage4	Stage5	Stage6	
	EVENT-BASED FEATURES							
1	Flu-Relatedness	✓	✓	✓	X	✓	✓	83%
2	Fire-Relatedness	X	X	✓	X	X	✓	33%
3	Cyber-Relatedness	X	✓	✓	X	X	✓	50%
4	Weaponization	X	X	✓	X	X	X	17%
6	OPINION-BASED FEATURES							
7	Positivity	X	X	X	X	X	X	0%
8	Negativity	X	X	X	X	X	X	0%
9	Neutrality	X	X	X	X	X	X	0%
10	Entropy	X	X	✓	X	X	X	17%
11	Sentiment	X	X	✓	X	X	X	17%
	EMOTION-BASED FEATURES							
12	Anticipation	X	X	✓	X	X	X	17%
13	Surprise	X	✓	✓	X	X	X	33%
14	Joy	X	X	✓	X	X	✓	33%
15	Trust	✓	X	✓	X	X	X	33%
16	Fear	✓	X	✓	✓	✓	X	67%
17	Sadness	X	✓	X	X	X	X	17%
18	Disgust	X	X	✓	X	X	X	17%
	OTHER							
19	Chatroom Congestion	✓	X	X	X	X	✓	33%
	TOTAL	**25%**	**25%**	**93%**	**6%**	**12%**	**31%**	**TOTAL**

Table 6-2: Predictive Features on the Social Dimension

The table above shows the predictive importance of each feature on the social dimension at each stage of the kill-chain. As stage 3 of the kill chain is seen by most cyber analysts as the network blind-spot, it is first observed that the social dimension is significantly important in predicting stage phase three of the kill-chain; the weaponization phase.

Furthermore, event-based features and emotion-based features are seen to be significant predictive indicators across all stages of the kill-chain. Significant events of importance within the time frame under observation such as the epidemic flu-spread and truck accident (from the scenario design), are directly linked to the propagation of events across the kill-chain. Event-based features are also seen to be not significantly predictive of events at the delivery stage of kill-chain (Stage 4 of the experiment). In general, features on the social dimension are observed to be not significantly important for predicting events at the delivery, exploitation and installation phase of the kill-chain. These results support two important notions; i) the structure of the kill-chain and ii) the manner of execution of each event in the kill-chain.

6.4.2 Predictive Features on the Physical Dimension

The physical dimension in the experimental design comprises of features on the real-world layer and the network layer. Several features from these layers were observed to be significantly important in the entangled cyberspace framework. The scenario and experimental design use network layer features from 3 sources (Packet flow log files, Intrusion detection log files and firewall log files). These feature sources record the flow of information in and out of the network, the intrusions attempted on the operating network and the type of connections made to the network. On the real-world layer, the average wind speed provides weather data observations for the metropolitan city in the scenario design. This weather data is integrated with microblogging feeds from the same population on the social dimension. This research aggregates network data measurements at equally time-spaced intervals to create aggregated time signals that characterise the operating network activities.

SN	FEATURE	STAGE1	STAGE2	STAGE3	STAGE4	STAGE5	STAGE6	
	PACKET FLOW FEATURES							
1	**Total Number of Packets**	X	X	X	X	✓	X	16.7%
2	**Mean Epoch Time**	X	X	X	X	✓	X	16.7%
3	**Mean Ip Length**	X	X	X	X	✓	X	16.7%
4	**Mean Transmitted Bytes**	X	X	✓	X	✓	X	33.3%
5	**Total number of UDP Connections**	X	✓	X	X	X	X	16.7%
6	**Total Number of connections**	X	X	X	✓	✓	X	33.3%

	to Database servers							
7	**Total Number of ICMP Packets**	X	X	X	X	X	✓	16.7%
	FIREWALL FEATURES							
8	**Total number of inbound connections**	X	X	X	✓	✓	✓	50%
9	**Total number of outbound connections**	X	X	X	✓	✓	✓	50%
10	**Total number of connections built**	X	X	X	✓	X	X	16.7%
11	**Total number of connections torn down**	X	X	X	✓	X	X	16.7%
12	**Total Number of services**	X	X	X	X	✓	X	16.7%
13	**Total Number of Operation**	X	X	X	X	X	X	0%
14	**Average Syslog Priority Level**	X	X	X	X	X	X	0%
	INTRUSION DETECTION FEATURES							
15	**Total number of IDS alerts**	X	✓	X	X	X	X	16.7%
16	**Total Warning Alerts**	X	✓	X	X	X	X	16.7%
17	**Average level of intrusion priority**	X	X	X	X	X	X	0%
18	**DOS Alerts**	X	X	X	X	X	X	0%
19	**Port Scan**	X	X	X	X	X	X	0%

	Alerts							
20	**Ping sweep Alerts**	X	X	X	X	X	X	0%
21	**Spp_frag3**	X	✓	X	X	X	X	16.7%
	DERIVED FEATURES							
22	**Recon**	X	X	X	X	X	✓	16.7%
23	**Connection Fail Ratio**	X	✓	X	X	X	X	16.7%
24	**Injection**	X	X	X	✓	X	X	16.7%
25	**Botnets**	X	X	X	X	✓	✓	33.3%
26	**DOS**						✓	16.7%
	TOTAL	**0%**	**19%**	**4%**	**23%**	**39%**	**23%**	

Table 6-3: Predictive Features on The Physical Dimension

The table above shows the predictive importance of each feature on the physical dimension at each stage of the kill-chain. The predictive importance of features on the physical dimension are observed to be sparsely spread across the phases of the kill-chain. The features are also observed to be more useful in predicting events at stages 5 & 6 (C&C and Attack) of the kill-chain. At this stage of the kill-chain, it is assumed that the adversary has certain control over the victim's network.

In general, the physical dimension, therefore, presents 4 major categories of features that are shown to be important in predicting events on the cyber-attack kill chain. These features include packet flow data, intrusion detection log data, derived features on the network layer and real-world observations particular to certain events of interest.

6.4.3 Predictive Features on the Economic Dimension

The features on the economic dimension as used in this research pertain to establishing the base for antecedents to cyber-attacks. The economic dimension, as it applies to this research, is theorized to contain features from political, cultural and financial real-world events. The scenario development incorporated features from digital financial markets as a representation of the economic dimension. The specific features used in the experimental design includes the average volume of the company's shares of a period of time, the weighted price index, the opening, closing, high and low prices of the shares of a period of time. The weighted price index is used as an endogenous variable in phase 1 of the experiment. The results from phase one of the experiments show a certain level of co-integration between activities on the economic dimension and activities on the social dimension. The average wind speed from the physical dimension was also seen to be a useful exogenous feature in predicting the weighted prices of the company's shares.

One further discussed limitations of this research is the restriction of the representation of the 'economic' layer to financial data. As earlier stated in the literature review, data on this dimension cover a range of SPEC factors that should ideally be grouped and tested. The addition of this layer serves as a proof of concept.

6.5 THE ENTANGLED CYBERSPACE FRAMEWORK

The proposed framework in chapter 2 of this report suggests an integration of two existing theories: the multi-dimensional cyberspace (Barnett, 2014) and the traditional kill-chain model (Hutchins, 2011). In addition to the traditional kill-chain model, this research incorporates the claims of social, political and economic effects on cyber-incidents by introducing a pre-stage to the traditional kill-

chain model; the antecedent phase. The results of the experiment show an interdependent relationship between events that characterize this phase and events from other stages of the kill-chain. Furthermore, the economic dimension is also incorporated into the entangled cyberspace framework to characterize the antecedent phase of the kill-chain.

Therefore, each phase of the kill-chain corresponds to a single dimension of cyberspace based on time observations of events. Events are therefore represented as time signals, where each phase can be characterized by multiple time signals representing multiple events on that phase. Multiple time signals are analysed using structural models for multiple time series analysis (Vector Auto-regressive models) as discussed in the literature review. The prediction accuracy and error measured at each stage of the experiment represents the usefulness of selected events in characterizing the defined stage of the kill-chain on the intended dimension. This experiment is carried out across the 5 phases of the denial of service kill-chain (Konikoff, Harris and Petersen, 2013) and predictive features are identified to characterize corresponding stages of the kill-chain.

The experiment testing the scenario was done in 6 stages which examines the events at the 6 stages of the attack scenario (including the antecedent stage) in this study. The research aims to provide empirical evidence for the existence of the entangled cyberspace using an integrated framework that detects early warning signs of cyber-incidents using a collection of webbed data. The theoretical framework of the entangled cyberspace is therefore practically implemented with the research approach. The research approach used identifies and integrates multiple sources of the evidence in the form of time signals, from various dimensions of cyberspace and analyses them using a multi-dimensional phased approach. The phases of the framework are assumed to be equivalent to the phases of the attack kill-chain under consideration (in this experiment, a denial of service attack). This research, therefore, provides empirical evidence for the proposed framework put forward in chapter 2 of this report. The proposed framework is an integration of the multi-dimensional cyberspace (Barnett, 2014), the traditional cyber-attack kill-chain (Hutchins, 2011) and Vector autoregressive models (Sims, 1980).

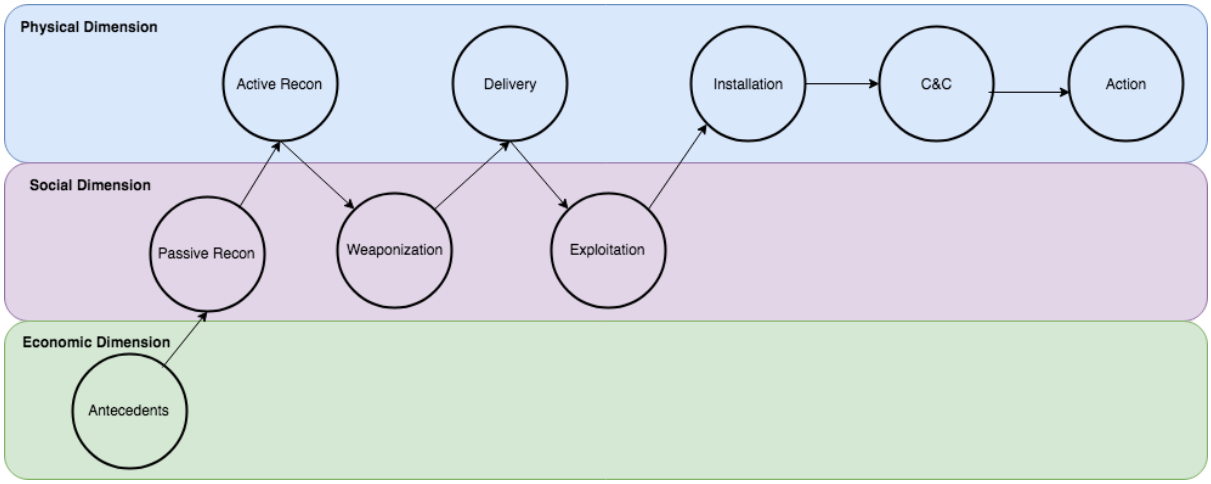

Figure 6-4: Scenario-Based Entangled Cyberspace Framework

The diagram above depicts a representation of the entangled cyberspace framework as applied to a hypothetical scenario. Each circle represents a phase in the cyber-attack kill-chain while the directionality of the arrows depicts the phased movements of events, vertically across the information space and horizontally across the kill-chain. The analytical framework as depicted in figure 5-2 is applied at each phase of the kill-chain on the corresponding dimension of representative evidence source. The entangled cyberspace, in theory, is the fusion of three conceptual foundations: a) A multi-dimensional characterization of cyberspace, b) A sequential phased model for perpetrating cyber-

attacks and c) A structural model for integrating and simultaneously analysing multiple sources of evidence. The entangled cyberspace is established in this research as a seamless integration of evidence sources useful for a certain domain of analysis. The domain of analysis under consideration in this research is the cyber domain, and the experiment is developed in the context of certain cyber-attack on a target network. The theoretical development for the research proposed framework is also developed in the context of a cyber-attack being perpetrated in cyberspace.

The entangled cyberspace starts with the characterization of cyberspace into different dimensions of interest. Various researchers as identified in the literature review have provided a wide range of cyberspace characterizations. After a suitable characterization of cyberspace has been identified, the framework identifies evidence sources across these characterized dimensions and integrates them in the context of an identified analytical domain. It assumes that all entities relevant to analytical tasks exist on some identified dimension of cyberspace.
These identified events are also characteristic of the stages of the traditional kill-chain model and used as the basis for developing the theoretical framework for pre-empting cyber-attacks. Given the domain of analysis (a cyber-attack), this research expands the idea of a seamless integration of evidence source across multiple dimensions of cyberspace by incorporating the steps involved in the traditional method for carrying out cyber-attacks. The stages of the traditional kill-chain model are split across the identified dimensions of cyberspace to create a multi-dimensional entanglement of events, characteristic of the kill-chain.
In the hypothesized scenario of this research, the task is predicting a cyber-attack which is represented as the last stage of the traditional cyber-attack kill-chain. The framework in this research tests the usefulness of prior events on each dimension of cyberspace in predicting successive stages of the kill-chain. Each prior stage is assumed to have happened before the start of the next stage of the kill-chain. The structural model propagates errors across the stages of the kill-chain, and only features that actively reduce prediction errors or increase prediction accuracy are used in predicting successive stages of the kill-chain. The structural model used is the vector autoregressive model due to its ability to extract relationships from multiple time series regression analysis simultaneously. The prediction accuracy and error measurements indicate the ability of events at one stage of the kill chain to pre-empt the events at the next phase kill-chain.

6.6 SUMMARY OF RESEARCH FINDINGS

The experiment testing the scenario in this research was conducted in 6 stages which examines the events at the 6 stages of the attack scenario in this study. The research aim is to provide an integrated framework that identifies early warning signs of cyber-incidents given a collection of webbed data. The research approach used identifies and integrates multiple sources of evidence from various dimensions of cyberspace and analyses them using a multi-dimensional phased approach. The framework can be interpreted as consisting of M phases and N dimensions where the phases correspond to the stages of the observed attack kill-chain while the dimensions correspond to the number of identified data sources (i.e. data dimensions) in cyberspace. This research, therefore, puts forward a framework that is an integration of evidence-based scenarios from a multi-dimensional cyberspace and a traditional cyber-attack kill-chain model. A benchmarked scenario was used to test the implementation of the theoretical framework, and the results of the multi-dimensional phased experiment are documented in chapter 4 of this report. This chapter discusses the findings from the experiment as they meet the specified research objectives. The findings from each stage of the experiment are detailed below, and the interpretation of the meaning is also provided.

The first stage of the analysis presents a contextual timeline for events in the scenario design. This timeline provides a time spectrum on which the hypothesized link between events across the three dimensions is tested. This stage of analysis justifies the use of identified sources of evidence in representing events in cyberspace. The second stage of analysis in this chapter discusses the results of inter-dimensional and intra-dimensional correlation, co-integration and causation. These stages of the analysis establish the links between events on the various dimensions of cyberspace. The third stage of this chapter analysis identifies the features relevant for predicting the perpetration of a cyber-attack across the various dimensions of cyberspace. Lastly, the final stage of analysis in this chapter establishes the Entangled Cyberspace as a seamless integration of inter-connected events across the various dimensions of cyberspace. This concept of inter-dependence of events in cyberspace has been tested and proven in this research using traditional cyber-attack monitoring frameworks and models of characterizations in cyberspace.

This method of analysis follows a strategy to fully justify the existence of the entangled cyberspace by individually meeting the requirements for each research objective stated. The analysis in this chapter was conducted in such a way that the contributions of the two frameworks used in developing the research artefacts are critically analysed. Firstly, the research justifies the sources of evidence as suggested by Alex Barnett, Clark and Klimburg in the Multi-Dimensional Cyberspace. Secondly, the sources of evidence that represent the various stages of Lockheed Martin's cyber-attack kill-chain are also evaluated and justified. This integrated analysis helps to justify the entanglements of events in cyberspace thoroughly.

6.7 CONCLUSION

This chapter has discussed various issues based on the research finding from the experiment and the existing literature on the implementation of analytical strategies to cyber defence. The discussions presented are based on thorough analysis of the benchmark data provided and extraction of hidden entanglements using the proposed framework. The research findings indicate that entanglements of events in cyberspace can be analysed using structural models for simultaneous analysis of multiple events. More so, the research findings also highlight specific predictive features of cyber incidents within each identified dimension of cyberspace. Each feature is evaluated based on its usefulness in predicting events at certain stages of the kill-chain.

The analysis of the experiment shows significant inter-dimensional relationships between feature types at various stages of the kill-chain. The antecedent phase is assumed to occur before the start of the kill-chain and assumingly kick starts the events of the kill-chain. The experimental timeline of events in phase one shows a chronological sequence of events. The chronological arrangement of events in the experimental design justifies a phased approach to analysing the causal effects of events between phases and across the dimensions of cyberspace. The phases of the experiment identify co-integrating events that are assumed to occur on each dimension with links to other dimensions of cyberspace. The results at the phased experimentation show that there are inter-dimensional and intra-dimensional correlative and co-integrating relationships.

The social dimension stands out with the most useful predictive features for events on other dimensions of cyberspace. The social dimension provides measures for the level of emotions shown by users towards specific events happening on the economic dimension. Events on the social dimension, across various phases of the experiment, are also proven to be predictive of events on another dimension.

Additionally, the economic dimension is proven to be useful in identifying the offset a kill-chain based on political, cultural and economic events. It is important to note that the economic dimension here represents political and economic factors that are assumed to influence the propagation of cyber-

attacks in cyberspace. The evidence identified on the economic dimension of cyberspace was shown to characterize the first phase of the modified cyber-attack kill-chain, the antecedent phase. This phase, therefore, characterizes events that trigger the start of a cyber-attack and the evidence to this fact is the predictive usefulness of these indicators as measured in phase one of the experiment.

The discussions at each phase of the experiments lead up to establishing the basis for the use of the entangled cyberspace theory in predicting cyber incidents. The analysis at each phase of the experiment presents the process and features to consider while including evidence sources in the framework. Additionally, co-integrating relationships were tested at each phase of the experiments to determine the extent to which features on at each phase of the kill-chain are inter-dependent.

Practically, the entangled cyberspace approach will contribute to an integrated implementation of proactive cyber defence techniques that can keep defenders one step ahead of their adversaries. The entangled cyberspace framework is based on a time-space categorization of events that are assumed to characterize stages of the traditional kill-chain which leads up to a cyber-attack. The next chapter presents the conclusion, contribution and further research for this study.

7 CHAPTER 7: CONCLUSION

7.1 INTRODUCTION

This chapter presents the conclusions of from the research findings for the entire study that was carried out to achieve the research aims and objectives. This research provides an understanding of the usefulness of certain entanglements of events in cyberspace on the propagation of cyber incidents. The researcher attempts to develop a theoretical framework to holistically incorporate entanglements of webbed data across the various dimensions of cyberspace in pre-empting cyber-attacks. Prediction of cyber-attacks is done with a structural phased approach, where phases correspond to the stages of the traditional cyber-attack kill-chain. This chapter summarizes the structure the structure and conduct of this study by reviewing the outcomes of deliverables of the thesis and then analysing the findings. Finally, this chapter presents the research contributions; contributions to theory and contributions to practice. The research limitations are considered and discusses suggestions for future research.

7.2 RESEARCH SUMMARY

Chapter one introduced the research agenda and highlighted the research motivation, research aims, research objectives, research strategy and intended contributions to theory and practice. **Chapter 1** emphasizes the importance of implementing proactive cyber defence strategies in modern organization infrastructure to curb the growing sophistication of cyber-attack methods used by adversaries. The increased global use of inter-connected devices and the increase in openness to information has created a sophisticated problem of the 'defence-attack' loop. Organizations would benefit from implementing proactive defence strategies that keep them ahead of the adversaries. Although organization are adopting are adopting a certain level of proactive cyber defence that includes monitoring, detection and mitigation techniques, a strategy that involves identifying early warning signs of cyber incidents an therefore pre-empting them, to achieve a superior level of cyber awareness over the adversary.

Chapter 2 expatiates on the theoretical rationale for the thesis by introducing and reviewing existing frameworks, techniques and methods for implementing proactive cyber situational awareness in organisations. The literature on the practical implementation of methods incorporated into the theoretical framework proposed in this research is also reviewed. The literature covered, different aspects of the entangled cyberspace theory including characterisation of cyberspace, evidence validation, data fusion, multivariate structural time series analysis and the cyber-attack kill-chain. The gaps in the literature were also identified, therefore highlighting the novelty of this study. The literature review highlighted a lack of existing frameworks to integrate multiple sources of evidence and methods for simultaneously analysing multiple entangled signals from events in cyberspace. Finally, the conceptual foundations for the theoretical framework were introduced in chapter 2, and the research hypotheses were also formulated. The conceptual foundation proposed incorporated three main individual concepts:

a) A multi-dimensional characterisation of cyberspace
b) A sequential phased model, the cyber-attack kill-chain, for perpetrating cyber-attacks
c) A structural model, Vector Autoregressive Models for integrating and simultaneously analysing multiple sources of evidence.

Chapter 3 presented the theoretical and conceptual foundations of this research. It critically discussed how the entangled cyberspace framework is developed using a combination of existing theories

discussed in the literature. Thus, addressing how the existing theories/models are used to ground the entangled cyberspace framework

Chapter 4 highlights the analysis of the research methodology used in this study to examine the implementation of the entangled framework in pre-empting cyber-attacks. This chapter discusses in detail the research approach briefly highlighted in chapter 1.

Additionally, the research assumptions were clearly stated and the effects that they may have on the results of the experiment. For example, the theoretical framework assumes all events in cyberspace occur on a pre-identified dimension of cyberspace and can be measured on a time-space continuum.

Additionally, the framework assumes all events are represented as time signals. Therefore, multiple sources must be measured on near similar timelines to fit into the multivariate model. Lastly, the multi-variate structural VAR assumes the same lag order for all time signals in the model. The limitations of a static lag-length selection were also discussed in the literature.

Chapter 4 also develops an acting scenario for testing the theoretical framework. This scenario testing is done using an experimental design which outlines how the experiment was designed around the scenario to test the validity of the theoretical framework developed in chapter 2. The multi-dimensional model used was introduced and vector autoregressive models with co-integrating time signals for events at each stage of the kill-chain was developed. Here, the researcher develops the experimental hypothesis to meet the set research objectives in chapter 1. The research hypothesis were set up to:

i) Identify the features that predict the occurrence of a cyber-attack on the network layer across the entangled information space.
ii) Test for correlation and co-integration of events between and within the identified dimensions of the entangled information space.
iii) Establish a predictive link between these activities between and within the identified dimensions of the entangled information space.
iv) Explain the nature of entanglements between these identified events and usefulness in pre-empting cyber-attacks.

Chapter 5 presents a practical implementation of the experiment based on events in the developed scenario. Time signals at each stage of the kill-chain, on their corresponding dimensions were analysed and the accuracy of models were evaluated by calculating the prediction errors at each stage. Each stage of the experiments represents events at each stage of the kill-chain, and it is assumed that these events occur in a multi-dimensional space of both kill-chain events and cyberspace dimensions respectively. In this chapter, the objectives of the research were met as follows:

i. Algorithms in cyber situational awareness, time series analysis, natural language processing and model evaluations are identified as useful methods and tools to be implemented in the integrated approach for pre-empting cyber-attacks. These methods are seen to capture the nature of entanglements between activities in the entangled information space.
ii. The dimensions of the entangled information space were proven as the economic dimension which consists of the market and media layers, the physical dimension which consists of the network and real-world layers and the social dimension which consists of the persona and cyber persona. The entanglements of events on the economic dimension and results from phase 1 of the experiment supports previous research by (Gandhi *et al.*, 2011b) in proving the effects of social, cultural, political and economic events on the proliferation of cyber-attacks.
iii. Features such as network packet flow features, text emotion, text sentiments and economic and financial indicators such as share prices are seen to be useful in developing models that predict the occurrence of cyber-attack on the network layer. An integration of these features

are also seen to explain the nature of entanglements between events across the dimensions of cyberspace.

Chapter 6 gives detailed discussions of the results from each stage of the experiment. Additionally, this chapter aligns the experiment results with the research aims and objectives by answering the research questions put forward in chapter 1. Chapter 5 also presents analytical discussions of the results from data analysis in view of the research literature. Chapter 5 extracts the classes of features indicative of cyber-incidents from an analytical perspective and identifies inter-dependencies between events across the various dimensions of cyberspace used in this research.

Chapter 7 presents an overview of previous chapters as well as the research contributions to theory, methodology and practice. This chapter also summarises the findings of the research and outlines the research limitations. Finally, suggestions for future research and research continuation are put ford with the domain of analysis.

In summary, this thesis presents an integrated approach for identifying entanglements of events in the various information layers of cyberspace. This thesis identified the physical and social dimension as used by previous authors in the literature and implemented the economic dimension as suggested by (Gandhi *et al.*, 2011a). The thesis put forward an approach that integrates methods of data fusion, data processing and analysis to untangle the complex dynamics of events in cyberspace across the information layers. After implementing the integrated approach to a set of representative entangled datasets, the findings of the thesis clearly show the possibilities of successfully pre-empting cyber-attacks by implementing the set of approaches suggested. The findings of this research also show the activities of the cyber-attack kill-chain hopping across dimensions these activities of the can be identified through intra-dimensional analysis. This confirmed integrated approach is presented in the figure 7.1 below. The figure depicts an integrated approach for identifying evidence sources to test for entanglements in across the information space. It is also a contextual summary of the experimental approach confirmed by the researcher for pin-pointing entanglements in the information space.

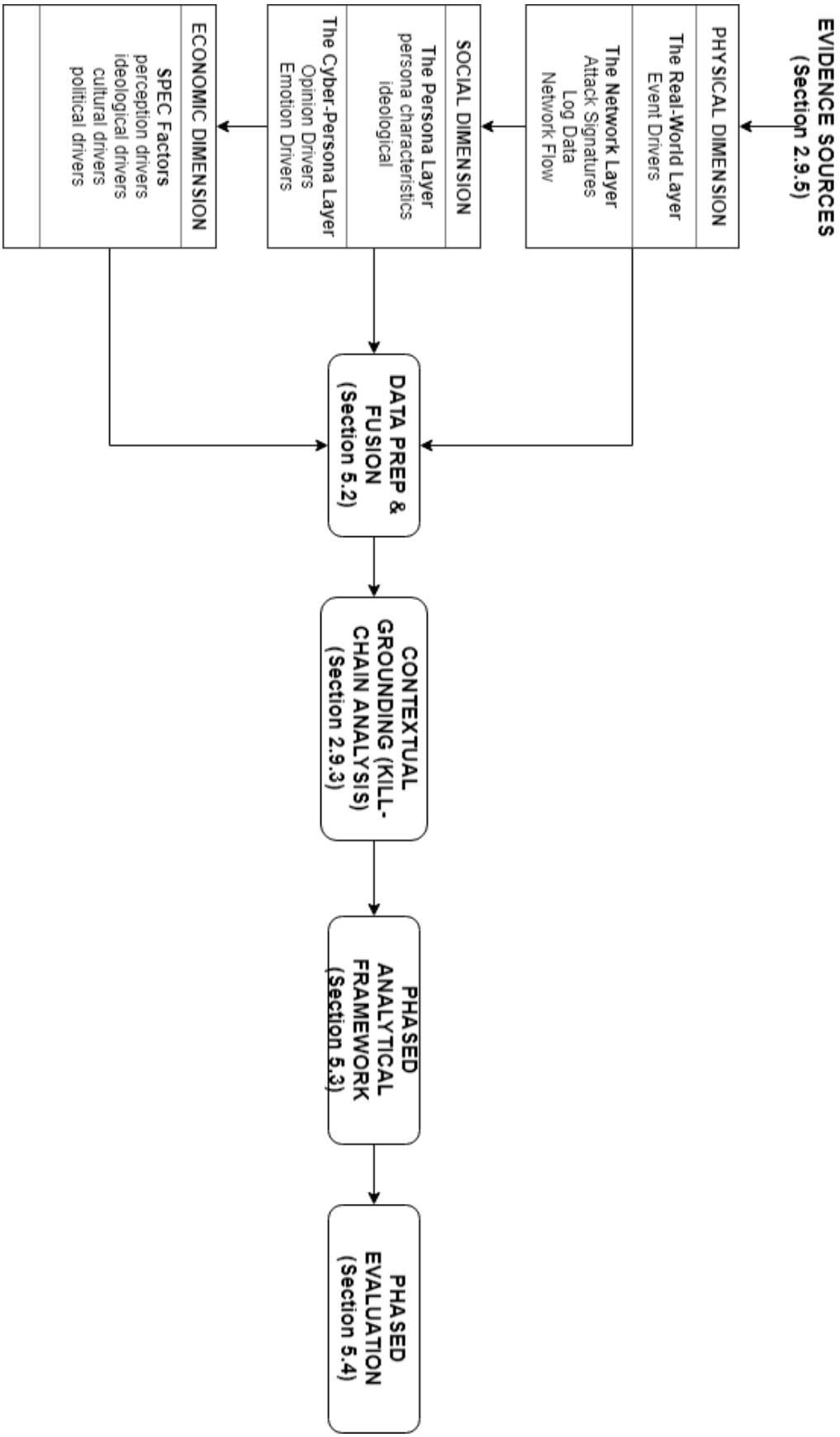

Figure 7-1: Integrated Approach for Untangling events in cyberspace

7.3 RESEARCH CONTRIBUTIONS

This section discusses the contributions of this research to the body of knowledge in the area of cyber defence. The study offers contributions in two categories: a) contributions to theory and b) contributions to practice. These contributions are further discussed in the following subsections.

7.3.1 Contributions to Theory

In chapter 2 of this study, the research identified the limited existence of a proactive cyber defence strategy capable of analysing the entanglements of events that affect cyber incidents. It was recognized that a multi-dimensional analytical model that incorporates multiple sources of evidence from various sources to simultaneously pre-empt cyber-attacks had not been considered in previous studies. Additionally, it was also identified that previous studies had limited analysis to a single dimension of cyberspace usually the physical dimension or the social dimension. Moreover, an analytical understanding of the antecedents to cyber-incidents had not been thoroughly visited in literature. The theoretical contributions of this research are therefore listed below:

a) ***Enhancing the existing cyber defence strategies, by providing a framework for identifying sources of evidence that are useful to pre-empting cyber incidents (Section 3.2).***
b) ***An approach for identifying features that are characteristic of various events across the various dimensions of the information space.***
c) ***A conceptual understanding of the features that are relevant for predicting cyber-incidents on identified dimensions of cyberspace (Section 6.5).***
d) ***A structural approach for co-integration and causal testing of events across cyberspace.***
e) ***A conceptual addition of the effects of Social-Political-Economic and Cultural (SPEC) on the proliferation of cyber-incidents.***
f) ***An introduction of a Social-Political-Economic and Cultural (SPEC) kill-chain for modelling the effects of SPEC factors on the proliferation of a cyber-attack (Section 6.6).***

This study contributes to IS literature as it introduces a multi-dimensional analytical approach to other major parts of IS research in cyber defence. Significant studies in cyber defence analysis revolved around a single linear analysis of data from single sources of evidence. These studies were limited in their ability to understand the dynamics of webbed data related to cyber-incidents. Additionally, only few studies (Hahn *et al.*, 2015) have provided a phased analysis of the traditional kill-chain. For example, (Hahn *et al.*, 2015) provides a phased analysis of the kill-chain although this analysis is limited to include only evidences from the network layer. Additionally, (Ning *et al.*, 2016) provides the theoretical background for the implementation of a cyber-physical-social-hyper thinking space that incorporates elements from the social, physical and information dimensions of cyberspace. They laid the foundation for the development of proactive cyber defence models by identifying elements that influence cyber events beyond the network layer. The entangled cyberspace framework put forward in this research is a multi-dimensional phased model that fills these inherent gaps in literature. The framework addresses the limited evidence base in the work of (Hahn *et al.*, 2015) and provides an empirical justification of certain parts of (Ning *et al.*, 2015) by incorporating structural models of analysis for multiple sources of evidence. These models are capable of simultaneously analysing multiple events and identifying hidden interconnections between events.

In addition to filling the analytical gaps of previous cyber defence models, this study contributes to IS literature on the information security management model (TM Forum, 2011) specifically, the monitoring, analysis and detection phases. The literature identifies gaps in current cyber defence analytical models about the sources of evidence that feed these models. Most of these evidence sources are mostly from a single dimension of cyberspace; the physical dimension specifically the

network layer or the social dimension. Moreover, on the social dimension, most cyber analytical techniques have been based solely on data from microblogging platforms such as Twitter. In addition to the works of (Hahn *et al.*, 2015; Hernández *et al.*, 2016),this research provides insights on identifying useful sources of evidence in cyberspace. As earlier stated, this research takes a multi-dimensional perspective by extending the scope of the evidence base for cyber analytics beyond the physical dimension; most especially the network layer.

Furthermore, this study builds on the works of (Klimburg and Mirtl, 2012; Barnett, 2014) by providing three similar characterisations of cyberspace based on certain functionalities of cyberspace in modern existence. These characterizations, in turn, serve as a potential evidence base for cyber defence predictive models. Therefore, the characterisation of cyberspace provided in this study serves to address the modern challenges of cyber defence in relation to the inter-relationships between cultural, political and economic dynamics in the real-world and the propagation of cyber-incidents in cyberspace. Additionally, this study shows that the use of a multi-dimensional evidence base integrated framework improves the quality of predictive models in cyber defence.

This study also largely contributes to cyber defence research in areas that seek to identify early warning indicators of cyber incidents. Although studies by (Kruegel *et al.*, 2003; Hernández *et al.*, 2016) provide useful models for feature identification for cyber threats, these features are usually mono-dimensional. The conceptual understanding of feature identification and extraction presented in this research harnesses features from multiple evidence-based using a phased approach. To this effect, this research offers a phased feature extraction conceptual model based on the phases of the traditional kill-chain. This research provides feature-based classifications for events at each phase of the kill-chain as they occur across the identified dimensions of cyberspace.

Furthermore, this research provides an overview of the predictive capacity of each dimension of cyberspace in predicting cyber incidents. Quantifiable measurements for user emotions and user sentiments are proven to significantly improve the quality of cyber predictive models on the social dimension.

Similarly, on the physical dimension, measurements for network alerts, network traffic and packet flow are identified as useful predictors of cyber-attacks. This corresponds with previous studies in cyber defence analytics such as (Kim, Paek and Oh, 2008; Konikoff, Harris and Petersen, 2013; Edkrantz, 2015; Stanfield and Stanfield, 2016). Lastly, the economic dimension as implemented in this research practically proves claims by (Gandhi *et al.*, 2011a) of an economic, cultural and political dimension to cyber incidents.

Finally, this study presents a new cyber-attack kill-chain perspective by the analysis of the traditional cyber-attack kill-chain to include social, political, economic and cultural factors presented in the literature review. This was achieved by providing a conceptual understanding of the antecedents of cyber-incidents and how they fit into the traditional cyber-attack kill-chain. The ‘antecedents’ in this research serves as a conceptual representation of SPEC factors and their effect on the propagation of cyber-incidents. Therefore, this research provides the first step for further incorporation of SPEC factors into the traditional cyber-attack kill-chain.

7.3.2 Contributions to Practice

The empirical findings from this study are also very useful to information security professionals such as security engineers, communications and network specialists, cyber risk analyst, cyber defence analysts, cyber data scientists etc. This study has practical influence for the implementation of the two major conceptual frameworks presented-(a) The entangled cyberspace and (b) The multi-level Kill Chain-with the ability to significantly improve the application of cyber defence strategies in

organisations. Thereby, enhancing the organisation's overall cyber situational awareness and improve the level of security by being ahead of the adversary. By implementing the phased framework provided in this research, cyber defenders can closely follow the propagation of each phase of the kill-chain and implement proper mitigation strategies to counteract adversaries' efforts.

Additionally, this study builds on and improves the information security management model by (TM Forum, 2011) as discussed in chapter 2 of this report. The monitoring, analysis and detection phases of the information security management model are critical to a successful proactive cyber defence strategy. This research, therefore, contributes to improving techniques and methods applied at these stages by providing certain insights to early warning indicators of cyber-indents beyond the physical dimension.

The results from the analysis in combination with gaps identified in the literature also show that organisations will benefit from building the capacity to handle a large analysis of structured and unstructured data from multiple data sources. This involves a large number of computational resources allocated to data identification, data extraction, data cleaning, data formatting, data fusion and data analysis. In addition to pre-defined data sources identified in the experimental design, this study provides a conceptual understanding of cyberspace characterisation which will help cybersecurity analyst determine what features are useful for prediction in certain cyber-attack scenarios.

Lastly, information security analysts can implement the frameworks put forward in this research to actively pre-empt the actions of adversaries in certain attack scenarios. The frameworks put forward- the entangled cyberspace framework and the SPEC cyber-attack kill-chain- are interdependent for pre-empting events in successive phases of the kill chain given events at the current stage of the kill-chain. Using the benchmark data, the entangled cyberspace framework is capable of untangling the entanglement of events in the context of the kill-chain. Therefore, security experts can implement such framework to understand the nature of events in their cyber operating space.

7.4 LIMITATIONS OF RESEARCH

This research attempted to advance theoretical and practical knowledge in cyber defence by developing a proactive, predictive model for cyber defence using a benchmark scenario. The research has achieved its aims and addressed the research objectives that were identified in chapter 1 of this report. However, it is important to identify the limitations of the study for possible further research.

Given that the kill-chain is a general model for the propagation of cyber-attacks, the framework is generalizable to all cyber-attacks assuming they adhere to the stages of the cyber-attack kill-chain. However, the experiment in this research is conducted in a single cyber-attack scenario - A denial of Service Attack on a Target Network. Therefore, the kill-chain used for the justification of the framework was restricted to kill-chain events in a DDOS attack.

The kill-chain limitation also raises the issue of extracted feature limitation, especially for features on the network layer. The features extracted in this research have been tested to improve prediction attempts in a DDOS attack scenario significantly. These features may, therefore, be limited to predicting events in the kill-chain of said cyber-attack scenario.

Additionally, the data used in the justification of the entangled cyberspace framework was based on benchmarked data. The benchmark data provided by (Cook *et al.*, 2011) is known to contain these entanglements of events and the framework put forward in this research proves capable of identifying and extracting these entanglements. However, the data used was pre-cleaned and structured which is usually not the case in real-life cyber analytical projects. Therefore, the framework is based on a pre-assumption of a certain structure and format to the data.

Furthermore, SPEC factors in this research are only represented with financial data. While this is sufficient to prove the practical application of the theory, the economic dimension goes further than the limited perspective presented in the scenario. The economic dimension serves to represent SPEC factors, as such, the scenario development used in this research lacks the robustness to encompass all theoretical aspects of SPEC effects on cyber-incidents.

The Entangled cyberspace framework is positioned to help put defenders ahead of the attackers, thereby increasing the overall level of cyber situational awareness in organisations. Another notable limitation is the fact that there is a need for empirical evidence to show if the level of cyber situational awareness is improved after implementing the strategies of the framework and also evaluate the impact on organisations' cyber defence strategy. There is also a need to measure the value added to organisations in terms of cyber-attack mitigations and recovery strategies.

Lastly, time and financial limitations hindered a rigorous longitudinal study to test the framework across multiple cyber-attack scenarios robustly.

7.5 FUTURE RESEARCH

This research has presented an integrated approach for tracking, monitoring and pre-empting cyber-attacks. Various methods for pin-pointing entanglements in cyberspace that inform prevention tactics in various cyber-attack scenarios. The approach presented in this research is extendible to various attack scenarios and open to alternative techniques for co-integration and causal testing. However, some of the limitations identified in section 7.4 provides some insights for useful contributions in research to improve the findings of this research. Firstly, the experimental designed provides the application of the approach in a denial of service attack scenario. Cyber-attack scenarios like malware infiltration, data exfiltration, SQL injections are of particular interest in cyber defence. Further research is encouraged to test and evaluate this approach in these alternative cyber-attack scenarios.

Technically, this research opens up the possibilities for simplifying co-integration testing between identified events across the various dimensions of cyberspace. In addition to the simplification of co-integration testing, this research also provides opportunities for in-depth causal testing between events on various dimensions. This advanced causal testing can include investigations into confirmatory analysis of factors that influence the propagation of cyber-attacks.

Additionally, since features at each stage of the kill-chain are not necessarily restricted to a single dimension as an evidence base, future research may extend this framework to include capabilities for robust intra-dimensional co-integration testing of features. This robust testing will address issues around automatic evidence identification and data fusion across multiple dimensions of cyberspace. Additionally, methods for automating the steps discussed in section 5.3 'The Analytical Framework' of this research would significantly improve the lag time in identification of antecedents to cyber-attacks, thereby, improving the quality of mitigation strategies.

Additionally, further research can address concrete justification of the effect of SPEC factors on the proliferation of cyber-attacks. The current approach justifies the existence of the economic dimension with access to simulated data representing SPEC features. Future research may incorporate features that characterise political and cultural events and test their predictive capacity in relation to cyber-incidents from a real-world perspective.

Lastly, due to time and financial constraints, the industry acceptance level and the willingness of cyber experts to implement the entangled cyberspace framework was not within the scope of this research. Future researchers will also need to work on measuring the technological acceptance level of the entangled cyberspace framework especially related to practical implementation in the industry.

Additionally, the willingness of cyber experts to implement the techniques in the entangled cyberspace framework may be a practical evaluation for the usefulness of the framework.

8 REFERENCES

Abdullahi, I., Arif, S. and Hassan, S. (2015) *Computational Intelligence in Information Systems*, *Advances in Intelligent Systems and Computing*. Edited by S. Phon-Amnuaisuk and T. W. Au. Springer International Publishing (Advances in Intelligent Systems and Computing). doi: 10.1007/978-3-319-13153-5.

Acock, A. (2012) 'What to Do About Missing Values', *APA handbook of research methods in psychology*, 3(2), pp. 27–50. doi: 10.1037/13621-002.

Adams, N. and Heard, N. (2014) *Data Analysis for Network Cyber-Security*. Imperial College Press. doi: 10.1142/p919.

Akaike, H., Clements, M. P. and Hendry, D. F. (1969) 'Fitting autoregressive models for prediction', *Annals of the Institute of Statistical Mathematics*, 21, pp. 243–247. Available at: http://onlinelibrary.wiley.com/doi/10.1002/9780470996430.ch1/summary.

Allahyari, M. *et al.* (2017) 'A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques'. Available at: http://arxiv.org/abs/1707.02919.

Almukaynizi, M. *et al.* (2017) 'Proactive Identification of Exploits in the Wild Through Vulnerability Mentions Online', *2017 International Conference on Cyber Conflict (CyCon U.S.)*, pp. 82–88. doi: 10.1109/CYCONUS.2017.8167501.

Alnabulsi, H., Islam, M. R. and Mamun, Q. (2014) 'Detecting SQL injection attacks using SNORT IDS', in *Asia-Pacific World Congress on Computer Science and Engineering, APWC on CSE 2014*. doi: 10.1109/APWCCSE.2014.7053873.

Anderson, T. . W. and Darling, D. . A. (1954) 'A Test of Goodness of Fit Author', 53(281), pp. 151–160.

Anscombe, F. J. and Glynn, W. J. (1983) 'Distribution of kurtosis statistic for normal statistics', *Biometrika*, 70, pp. 227–234.

Atzori, L. *et al.* (2012) 'The social internet of things (SIoT) - When social networks meet the internet of things: Concept, architecture and network characterization', *Computer Networks*, 56(16), pp. 3594–3608. doi: 10.1016/j.comnet.2012.07.010.

Axelrod, R. and Iliev, R. (2014) 'Timing of cyber conflict', *Proceedings of the National Academy of Sciences*, 111(4), pp. 1298–1303. doi: 10.1073/pnas.1322638111.

Baccianella, S., Esuli, A. and Sebastiani, F. (2010) 'SentiWordNet 3.0 : An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining SentiWordNet', *Analysis*, 10(January 2010), pp. 1–12. doi: 10.1.1.61.7217.

Baggili, I. and Breitinger, F. (2015) 'Data Sources for Advancing Cyber Forensics: What the Social World Has to Offer', *AAAI Spring Symposium*, pp. 2009–2012.

Bailey Lee, C., Roedel, C. and Silenok, E. (2003) 'Detection and Characterization of Port Scan Attacks', *Univeristy of California, Department of Computer Science and Engineering*, pp. 1–7. Available at: http://cseweb.ucsd.edu/~clbailey/PortScans.pdf.

Bar, A. *et al.* (2016) 'Identifying Attack Propagation Patterns in Honeypots Using Markov Chains Modeling and Complex Networks Analysis', *Proceedings - 2016 IEEE International Conference on Software Science, Technology and Engineering, SwSTE 2016*, pp. 28–36. doi: 10.1109/SWSTE.2016.13.

Barford, P. *et al.* (2002) 'A signal analysis of network traffic anomalies', *Proceedings of the second ACM SIGCOMM Workshop on Internet measurment - IMW '02*, p. 71. doi: 10.1145/637201.637210.

Barford, P. *et al.* (2010) 'Cyber SA: Situational awareness for cyber defense', *Advances in Information Security*, 46, pp. 3–13.

Barnett, A. (2014) '19th ICCRTS : C2 Agility : Lessons Learned from Research and Operations Paper 081 : Using causal models to manage the cyber threat to C2 agility : working with the benefit of hindsight', *International Command and Control Research and Technology Symposium*.

Barnett, A., Smith, S. and Whittington, R. (2014) 'Using causal models to manage the cyber threat to C2 agility : working with the benefit of hindsight', *19th ICCRTS: C2 Agility*.

Bartlett, J. (2015) *The Dark Net*. London: Windmill Books. doi: 10.1177/0740277515623750.

Bayes, T. (1763) 'An Essay Towards Solving a Problem in the Doctrines of Chances', *Philosophical Transactions*, 53(1764), pp. 370–418. doi: 10.1093/biomet/45.3-4.293.

Ben-Asher, N. and Gonzalez, C. (2015) 'Effects of cyber security knowledge on attack detection', *Computers in Human Behavior*. Elsevier Ltd, 48, pp. 51–61.

Bentz, C. (2016) 'The Word Entropy of Natural Languages a'.

Bera, A. K. and Jarque, C. M. (1981) 'Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Monte Carlo Evidence', *Economics Letters*, 7(4), pp. 313–318. doi: 10.1016/0165-1765(81)90035-5.

Berger, A. L., Pietra, S. A. Della and Pietra, V. J. Della (1996) 'A Maximum Entropy Approach to Natural Language Processing', *Computational Linguistics*, 22, pp. 39–71. doi: 10.3115/1075812.1075844.

Bhuyan, M. H., Bhattacharyya, D. K. and Kalita, J. K. (2011) 'Surveying port scans and their detection methodologies', *Computer Journal*. doi: 10.1093/comjnl/bxr035.

Bishop, C. M. (2006) *Pattern Recognition and Machine Learning*, *Pattern Recognition*. Edited by M. Jordan, J. Kleinberg, and B. Schölkopf. Springer (Information science and statistics). doi: 10.1117/1.2819119.

Bishop, P., Hines, A. and Collins, T. (2007) *The current state of scenario development: an overview of techniques*, *Foresight*. doi: 10.1108/14636680710727516.

Blumenstock, J. E. (2008) 'Size matters: word count as a measure of quality on wikipedia', *Proceedings of the 17th International Conference on World Wide Web*, pp. 1095–1096. doi: 10.1145/1367497.1367673.

Bollen, J. and Mao, H. (2011) 'Twitter Mood Predicts The Stock Market-ppt', *Journal of Computational Science*, 2(1), pp. 1–8. Available at:

http://www.technologyreview.com/view/421251/twitter-mood-predicts-the-stock-market/.

Bonev, B., Escolano, F. and Cazorla, M. (2008) 'Feature selection, mutual information, and the classification of high-dimensional patterns: Applications to image classification and microarray data analysis', *Pattern Analysis and Applications*, 11(3–4), pp. 309–319. doi: 10.1007/s10044-008-0107-0.

Borja, M. (2013) 'Instant Traffic Analysis with Tshark How-To', in. Packt Publishing Ltd.

Börjeson, L. *et al.* (2006) 'Scenario types and techniques: Towards a user's guide', *Futures*, 38(7), pp. 723–739. doi: 10.1016/j.futures.2005.12.002.

Bou-Harb, E., Debbabi, M. and Assi, C. (2015) 'A time series approach for inferring orchestrated probing campaigns by analyzing darknet traffic', *Proceedings - 10th International Conference on Availability, Reliability and Security, ARES 2015*, pp. 180–185. doi: 10.1109/ARES.2015.9.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (2008) *Time Series Analysis: Forecasting and Control*, *Wiley*. John Wiley & Sons, Inc. doi: 10.1016/j.ijforecast.2004.02.001.

Braun, P. A. and Mittnik, S. (1993) 'Misspecifications in vector autoregressions and their effects on impulse responses and variance decompositions', *Journal of Econometrics*, 59(3), pp. 319–341. doi: 10.1016/0304-4076(93)90029-5.

Breier, J. and Brani\textbackslashvsová, J. (2015) 'Anomaly Detection from Log Files Using Data Mining Techniques', *Information Science and Applications*, 339, pp. 449–457. Available at: http://dx.doi.org/10.1007/978-3-662-46578-3_53.

Bronk, C. and Tikk-Ringas, E. (2013) 'The Cyber Attack on Saudi Aramco', *Survival*, 55(2), pp. 81–96. doi: 10.1080/00396338.2013.784468.

Burnap, P. and Williams, M. L. (2016) 'Us and them: identifying cyber hate on Twitter across multiple protected characteristics', *EPJ Data Science*. Burnap and Williams, 5(1). doi: 10.1140/epjds/s13688-016-0072-6.

Calefato, F., Lanubile, F. and Novielli, N. (2017) 'EmoTxt: A Toolkit for Emotion Recognition from Text', 1(July), pp. 3–4. doi: 10.1109/ACIIW.2017.8272591.

Cambria, E. *et al.* (2013) 'New Avenues in Opinion Mining and Sentiment Analysis', *IEEE Intelligent Systems*, 28(2), pp. 15–21. doi: 10.1109/MIS.2013.30.

Canright, G. S. and Eng??-Monsen, K. (2008) 'Some relevant aspects of network analysis and graph theory', in *Handbook of Network and System Administration*. doi: 10.1016/B978-044452198-9.50017-3.

Cavusoglu, H., Mishra, B. and Raghunathan, S. (2004) 'The Effect of Internet Security Breach Announcements on Market Value: Capital Market Reactions for Breached Firms and Internet Security Developers', *International Journal of Electronic Commerce*, 9(1), pp. 69–104. doi: 10.1.1.85.3407.

cdc.gov (2016) *Flu Symptoms & Complications*.

Chakraborty, S. . b *et al.* (2016) 'Predicting socio-economic indicators using news events', in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

doi: 10.1145/2939672.2939817.

Chen, C. and Liu, L.-M. (1993) 'Joint Estimation of Model Parameters and Outlier Effects in Time Series', *Journal of the American Statistical Association*, 88(421), p. 284. doi: 10.2307/2290724.

Chen, Y. and Hwang, K. (2006) 'Collaborative detection and filtering of shrew DDoS attacks using spectral analysis', *Journal of Parallel and Distributed Computing*, 66(9), pp. 1137–1151. doi: 10.1016/j.jpdc.2006.04.007.

Cheung, Y. W. and La, K. S. (1995) 'Lag order and critical values of the augmented dickey-fuller test', *Journal of Business and Economic Statistics*, 13(3), pp. 277–280. doi: 10.1080/07350015.1995.10524601.

Cisco (2017) 'Identifying Incidents Using Firewall and Cisco IOS Router Syslog Events', pp. 1–8. Available at: http://www.cisco.com/web/about/security/intelligence/identify-incidents-via-syslog.html#2.

Clark, D. (2010) 'Characterizing Cyberspace: Past, Present, and Future', *ECIR Working Paper*, 2010(1984), pp. 1–18. Available at: https://projects.csail.mit.edu/ecir/wiki/images/7/77/Clark_Characterizing_cyberspace_1-2r.pdf.

Cleveland, W. S. and Sun, D. X. (1995) 'Internet Traffic Data', *Journal of the American Statistical Association*, 95(451), pp. 79–985.

CoinMarketCap (2018) *Bitcoin Historical Data*, *CoinMarketCap*. Available at: https://coinmarketcap.com/currencies/bitcoin/historical-data/.

Colneric, N. and Demsar, J. (2018) 'Emotion Recognition on Twitter: Comparative Study and Training a Unison Model', *IEEE Transactions on Affective Computing*, 3045(c). doi: 10.1109/TAFFC.2018.2807817.

Cook, K. *et al.* (2011) 'VAST 2011 Challenge : Cyber Security and Epidemic', *EEE VAST*, pp. 299–301.

Creswell, J. W. (2003) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches 3rd by John W. Creswell (2008) Paperback: Amazon.co.uk: Books*. 4th edn. SAGE Publications. doi: 10.4135/9781849208956.

D'Agostino, R. and Pearson, E. S. (1973) 'Tests for departure from results for the normality of b2 and √b1', *Biometrika*, 60(3), pp. 613–622. doi: 10.2307/2335012.

David, J. and Thomas, C. (2015) 'DDoS attack detection using fast entropy approach on flow-based network traffic', *Procedia Computer Science*. Elsevier Masson SAS, 50, pp. 30–36. doi: 10.1016/j.procs.2015.04.007.

Davis, G. B. (2000) 'Information Systems Conceptual Foundations: Looking Backward and Forward', *Organizational and Social Perspectives on Information Technology*, pp. 61–82. doi: 10.1007/978-0-387-35505-4_5.

DCDC (2015a) *Strategic Trends Programme Future Operating Environment 2035 First Edition*. Available at: https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&cad=rja&uact=8&ved=0ahUKEwj2p4-

E8a7WAhWrCMAKHeuWA4UQFghEMAM&url=http%3A%2F%2Fwww.defencesynergia.co.uk%2Fwp-content%2Fuploads%2F2016%2F08%2FDCDC-Future-Operating-Environment-out-to-2035-201.

DCDC (2015b) *The Future Character of Conflict*. Available at: https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&cad=rja&uact=8&ved=0ahUKEwj2p4-E8a7WAhWrCMAKHeuWA4UQFghOMAU&url=http%3A%2F%2Fwww.mod.uk%2FNR%2Frdonlyres%2FA05C6EB5-5E8F-4115-8CD6-7DCA3D5BA5C6%2F0%2FFCOCReadactedFinalWeb.pdf&usg=AFQjCNF_8.

Debatin, B. *et al.* (2009) 'Facebook and online privacy: Attitudes, behaviors, and unintended consequences', *Journal of Computer-Mediated Communication*, 15(1), pp. 83–108. doi: 10.1111/j.1083-6101.2009.01494.x.

Debole, F. and Sebastiani, F. (2003) 'Supervised term weighting for automated text categorization', *Proceedings of the 2003 ACM symposium on Applied computing - SAC '03*, p. 784. doi: 10.1145/952686.952688.

Development Concepts and Doctrine Centre (2013) 'Cyber Primer', p. 78.

Donier, J. and Bouchaud, J. P. (2015) 'Why do markets crash? Bitcoin data offers unprecedented insights', *PLoS ONE*, 10(10), pp. 23–25. doi: 10.1371/journal.pone.0139356.

Doupé, A. *et al.* (2011) 'Hit 'em Where it Hurts: A Live Security Exercise on Cyber Situational Awareness', *Proceedings of the 27th Annual Computer Security Applications Conference*, pp. 51–61.

Dua, S. and Du, X. (2013) *Data Mining and Machine Learning in Cybersecurity*, *Journal of Chemical Information and Modeling*. doi: 10.1017/CBO9781107415324.004.

Dwivedi, N. and Tripathi, A. (2015) 'Event Correlation for Intrusion Detection Systems', *2015 IEEE International Conference on Computational Intelligence & Communication Technology*. doi: 10.1109/CICT.2015.111.

Edkrantz, M. (2015) 'Predicting Exploit Likelihood for Cyber Vulnerabilities with Machine Learning', p. 57. Available at: http://studentarbeten.chalmers.se/publication/219658-predicting-exploit-likelihood-for-cyber-vulnerabilities-with-machine-learning%0Ahttp://publications.lib.chalmers.se/records/fulltext/219658/219658.pdf.

Ekman, P. (1992) 'An Argument for Basic Emotions', *Cognition and Emotion*, pp. 169–200. doi: 10.1080/02699939208411068.

Enders, W. (2004) *Applied Econometric Time Series*, *Technometrics*. doi: 10.1198/tech.2004.s813.

Enders, W. (2014) 'Chapter 2 - Stationary Time-Series Models', in *Applied Econometric Time Series*. John Wiley & Sons, Inc., p. 496.

Engel, G. (2014) *Deconstructing the Cyber Kill Chain*, *Dark Reading*.

Engle, R. F. *et al.* (2017) 'Co-Integration and Error Correction : Representation , Estimation , and Testing Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/1913236 REFERENCES Linked references are available on JSTOR for this article : You may need to ', 55(2), pp. 251–276.

Engle, R. F. and Yoo, B. S. (1987) 'Forecasting and testing in co-integrated systems', *Journal of Econometrics*, 35(1), pp. 143–159. doi: 10.1016/0304-4076(87)90085-6.

Eriksson, O., Olofsson, M. and Ekvall, T. (2003) 'How model-based systems analysis can be improved for waste management planning.', *Waste management & research : the journal of the International Solid Wastes and Public Cleansing Association, ISWA*, 21(6), pp. 488–500. doi: 10.1177/0734242X0302100602.

Eterovic, T. *et al.* (2014) 'Data mining meets network analysis: Traffic prediction models', *Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, (May), pp. 1479–1484. doi: 10.1109/MIPRO.2014.6859800.

Farhadi, H., Amirhaeri, M. and Khansari, M. (2011) 'Alert Correlation and Prediction Using Data Mining and HMM', *The ISC Int'l Journal of Information Security*, 3, pp. 77–101.

Fava, D. S., Byers, S. R. and Yang, S. J. (2008) 'Projecting cyberattacks through variable-length Markov models', *IEEE Transactions on Information Forensics and Security*, 3(3), pp. 359–369. doi: 10.1109/TIFS.2008.924605.

Feria, E. H. (2010) 'Latency-information theory', in *2010 IEEE Sarnoff Symposium*. IEEE, pp. 1–8. doi: 10.1109/SARNOF.2010.5469775.

Fischer, F. and Keim, D. a. (2014) 'NStreamAware: real-time visual analytics for data streams to enhance situational awareness', *Proceedings of the Eleventh Workshop on Visualization for Cyber Security, ACM*, pp. 66–72. Available at: http://dl.acm.org/citation.cfm?id=2671495.

Forum, T. M. and Reserved, A. R. (2011) 'TM Forum Security Management Model', pp. 1–30.

Freitas, A. (2013) *Data mining and knowledge discovery with evolutionary algorithms*. Springer Science \& Business Media.

Fu, Q. *et al.* (2009) 'Execution anomaly detection in distributed systems through unstructured log analysis', *Proceedings - IEEE International Conference on Data Mining, ICDM*, pp. 149–158. doi: 10.1109/ICDM.2009.60.

Fu, Y. (2016) 'Theory of interaction', *Theoretical Computer Science*. Elsevier B.V., 611, pp. 1–49. doi: 10.1016/j.tcs.2015.07.043.

Gabra, H. N. (2014) 'Classification of IDS Alerts with Data Mining Techniques', in *Proceedings - IEEE Military Communications Conference MILCOM*. Available at: http://arxiv.org/abs/1401.4872.

Galenko, A. *et al.* (2009) 'Simulating cointegrated time series', *Proceedings - Winter Simulation Conference*, (2002), pp. 483–493. doi: 10.1109/WSC.2009.5429356.

Gandhi, R. *et al.* (2011a) 'Dimensions of cyber-attacks: Cultural, social, economic, and political', *IEEE Technology and Society Magazine*, 30(1), pp. 28–38. doi: 10.1109/MTS.2011.940293.

Gandhi, R. *et al.* (2011b) 'Social, Political, Economic, and Cultural'.

García-Teodoro, P. *et al.* (2009) 'Anomaly-based network intrusion detection: Techniques, systems and challenges', *Computers and Security*, 28(1–2), pp. 18–28. doi: 10.1016/j.cose.2008.08.003.

Gehrke, J., Ginsparg, P. and Kleinberg, J. (2007) 'Overview of the 2003 KDD Cup', *ACM SIGKDD*

Explorations Newsletter, 5(2), p. 149. doi: 10.1145/980972.980992.

George Mason University (2010) 'Advanced Cyber Attack Modeling , Analysis , and Visualization', *Advance Cyber Attack Modelling, Analysis, and visualization*.

Gervais, M. (2012) 'Cyber Attacks and the Laws of War', *Berkeley Journal of International Law*, 30(2), pp. 525–579. doi: 10.2139/ssrn.1939615.

Gevrey, M., Dimopoulos, I. and Lek, S. (2003) 'Review and comparison of methods to study the contribution of variables in artificial neural network models', *Ecological Modelling*, 160(3), pp. 249–264. doi: 10.1016/S0304-3800(02)00257-0.

Ghasemi, A. and Zahediasl, S. (2012) 'Normality tests for statistical analysis: A guide for non-statisticians', *International Journal of Endocrinology and Metabolism*, 10(2), pp. 486–489. doi: 10.5812/ijem.3505.

Gotved, S. (2006) 'Time and space in cyber social reality', *New Media & Society*, 8(3), pp. 467–486. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-33746355993&partnerID=tZOtx3y1.

Granger, C. W. J. (1980) 'Testing for causality. A personal viewpoint', *Journal of Economic Dynamics and Control*, 2(C), pp. 329–352. doi: 10.1016/0165-1889(80)90069-X.

Granovetter, M. S. (1982) 'The strength of "weak ties": A network theory revised', *Social structure and network analysis*, 1(1983), pp. 105–130.

Grimmer, J. and Stewart, B. M. (2013) 'Text as data: The promise and pitfalls of automatic content analysis methods for political texts', *Political Analysis*, 21(3), pp. 267–297.

Haddi, E., Liu, X. and Shi, Y. (2013) 'The role of text pre-processing in sentiment analysis', in *Procedia Computer Science*, pp. 26–32.

Hahn, A. *et al.* (2015) 'A multi-layered and kill-chain based security analysis framework for cyber-physical systems', *International Journal of Critical Infrastructure Protection*.

Hair, J. F. (2014) *Multivariate Data Analysis*. Upper Saddle River, NJ: Pearson Prentice Hall. doi: 10.1016/j.ijpharm.2011.02.019.

Hall, M. a and Smith, L. a (1998) 'Feature subset selection: A correlation based filter approach', *Progress in Connectionist-Based Information Systems, Vols 1 and 2*, pp. 855–858.

Hamilton, J. D. (1994) 'Time Series Analysis', *Book*, p. xiv. doi: 10.2307/1270781.

Hannan, E. . and Quinn, B. . (2010) 'The Determination of the Order of an Autoregression', *The Annals of Statistics*, 38(3), pp. 205–247.

Harding, L. (2014) *The Snowden files*. London: Guardian Books.

Hathaway, O. A. *et al.* (2012) 'The law of cyber-attack', *California Law Review*, 100(4), pp. 817–885. doi: 10.15779/Z38CR6N.

HCIL (2013) *Visual Analytics Benchmark Repository*. Available at: https://www.cs.umd.edu/hcil/varepository/VAST Challenge 2013/challenges/MC2 - Situation

Awareness Display Design/.

Hern, A. and Gibbs, S. (2017) 'What is WannaCry ransomware and why is it attacking global computers? | Technology | The Guardian', *The Guardian*, May. Available at: https://www.theguardian.com/technology/2017/may/12/nhs-ransomware-cyber-attack-what-is-wanacrypt0r-20.

Hernández, A. *et al.* (2016) 'Security Attack Prediction Based on User Sentiment Analysis of Twitter Data', *Proceedings of the 2016 IEEE International Conference on Industrial Technology (ICIT)*, pp. 610–617. doi: 10.1109/ICIT.2016.7474819.

Hevner, A. R. *et al.* (2004) 'Design Science in Information Systems Research', *MIS Quarterly*, 28(1), pp. 75–105. doi: 10.2307/25148625.

Hong, Y. T. *et al.* (2003) 'Correlation between Indian Ocean summer monsoon and North Atlantic climate during the Holocene', *Earth and Planetary Science Letters*, 211(3–4), pp. 371–380. doi: 10.1016/S0012-821X(03)00207-3.

Hoover, K. D. and Perez, S. J. (1994) 'Post hoc ergo propter once more an evaluation of "does monetary policy matter?" in the spirit of James Tobin', *Journal of Monetary Economics*, 34(1), pp. 47–74. doi: 10.1016/0304-3932(94)01149-4.

Hornecker, E. and Buur, J. (2006) 'Getting a grip on tangible interaction: a framework on physical space and social interaction', *Chi*, pp. 437–446. doi: http://doi.acm.org/10.1145/1124772.1124838.

Hutchins, E. M. (2011) 'Cyber Kill Chain', *That Book*, pp. 80–106. Available at: http://www.lockheedmartin.com/us/what-we-do/information-technology/cybersecurity/tradecraft/cyber-kill-chain.html.

Hutchins, E. M., Cloppert, M. J. and Amin, R. M. (2011) 'Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains', *6th Annual International Conference on Information Warfare and Security*.

Hylleberg, S. *et al.* (1990) 'Seasonal integration and cointegration', *Journal of Econometrics*, 44(1–2), pp. 215–238. doi: 10.1016/0304-4076(90)90080-D.

Hyndman, R. J. and Koehler, A. B. (2006) 'Another look at measures of forecast accuracy', *International Journal of Forecasting*, 22(4), pp. 679–688. doi: 10.1016/j.ijforecast.2006.03.001.

Ioannou, G., Clewley, N. and Powell, G. (2013) 'A Markov Multi-Phase Transferable Belief Model : An Application for predicting Data Exfiltration', *16th International Conference on Information Fusion*, pp. 842–849.

Issues, R. (2012) 'Detecting packet injection : a guide to observing packet spoofing by ISPs', pp. 1–17.

Jackson, G. (2012) *Predicting Malicious Behaviour. Tools and Techniques for Ensuring Global Security*. 1st edn. John Wiley & Sons.

Jason, E. (2007) 'EMNLP-CoNLL 2007'.

Jin, S. J. S. and Yeung, D. S. (2004) 'A covariance analysis model for DDoS attack detection', *2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)*, 4(c), pp. 1882–1886. doi:

10.1109/ICC.2004.1312847.

Jin, X. and An, X. (2015) 'Global financial crisis and emerging stock market contagion: A volatility impulse response function approach', *Research in International Business and Finance*, 36, pp. 179–195. doi: 10.1016/j.ribaf.2015.09.019.

Johansen, S. (2000) 'Modelling of cointegration in the vector autoregressive model', *Economic Modelling*, 17, pp. 359–373. doi: 10.1016/S0264-9993(99)00043-7.

Johnson, W. and Bouchard, T. J. (2005) 'The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized', *Intelligence*, 33(4), pp. 393–416. doi: 10.1016/j.intell.2004.12.002.

Jumratjaroenvanit, A. and Teng-Amnuay, Y. (2008) 'Probability of attack based on system vulnerability life cycle', in *Proceedings of the International Symposium on Electronic Commerce and Security, ISECS 2008*. doi: 10.1109/ISECS.2008.212.

Jurafsky, D. and Martin, J. (2017) 'Vector Sematics', in *Speech and Language Processing*. 2nd edn. Prentice Hall, pp. 99–124.

Kadivar, M. and Buzan, B. G. (2014) 'Cyber-Attack Attributes', *Technology Innovation Management Review*, 4(11), pp. 22–28. doi: 1638204721.

Kaji, N. and Kitsuregawa, M. (2007) 'Building Lexicon for Sentiment Analysis from Massive Collection of HTML Documents.', *EMNLP-CoNLL*, 43(June), pp. 1075–1083.

Kao, C. N. *et al.* (2015) 'A predictive zero-day network defense using long-term port-scan recording', in *2015 IEEE Conference on Communications and NetworkSecurity, CNS 2015*. doi: 10.1109/CNS.2015.7346890.

Karatzogianni, A. (2008) *Cyber conflict and global politics*, *Cyber Conflict and Global Politics*. Routledge. doi: 10.4324/9780203890769.

Khan, F. H., Qamar, U. and Bashir, S. (2016) 'SentiMI: Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection', *Applied Soft Computing Journal*. Elsevier B.V., 39, pp. 140–153. doi: 10.1016/j.asoc.2015.11.016.

Khater, N. Al and Overill, R. E. (2015) 'Forensic Network Traffic Analysis', in *Proceedings of The Second International Conference on Digital Security and Forensics*.

Kick, J. (2014) 'Cyber Exercise Playbook', *Cyber Exercise Playbook*, 7013(November), pp. 1–40. Available at: https://www.mitre.org/sites/default/files/publications/pr_14-3929-cyber-exercise-playbook.pdf.

Kim, D., Paek, S. H. and Oh, H. S. (2008) 'A Hilbert-Huang transform approach for predicting cyber-attacks', *Journal of the Korean Statistical Society*, 37(3), pp. 277–283. doi: 10.1016/j.jkss.2008.02.006.

Kim, S. S. and Reddy, A. L. N. (2008) 'Statistical techniques for detecting traffic anomalies through packet header data', *IEEE/ACM Transactions on Networking*, 16(3), pp. 562–575. doi: 10.1109/TNET.2007.902685.

Kinable, J. (2008) 'Detection of network scan attacks using flow data', *9th Twente Student Conference on IT, 23th June*. Available at: http://www.utwente.nl/ewi/dacs/assignments/completed/bachelor/reports/2008-kinable.pdf.

Kipper, K. *et al.* (2006) 'Extending VerbNet with novel verb classes', *Proceedings of LREC*, 2006(2.2), p. 1.

Klimburg, A. (2011) 'Special Issue Cyberspace and Governance — A primer', (September).

Klimburg, A. and Mirtl, P. (2012) 'Cyberspace and Governance - A Primer', *Austrian Institute for International Affairs*, (September 2011). Available at: http://www.oiip.ac.at/fileadmin/Unterlagen/Dateien/Publikationen/Cyberspace_and_Governance_-_Working_Paper_65_2.pdf.

Konikoff, E., Harris, B. and Petersen, P. (2013) 'Breaking the DDoS Attack Chain', *Institute for Software Research*, (August). Available at: http://www.cmu.edu/mits/files/breaking-the-ddos-attack-chain.pdf.

Kruegel, C. *et al.* (2003) 'Bayesian event classification for intrusion detection', in *Proceedings - Annual Computer Security Applications Conference, ACSAC*. doi: 10.1109/CSAC.2003.1254306.

Kwak, H. *et al.* (2010) 'What is Twitter, a Social Network or a News Media?', *Network*, pp. 19–22. doi: 10.1145/1772690.1772751.

Lau, R. Y. K., Xia, Y. and Li, C. (2012) 'Social Media Analytics for Cyber Attack Forensic', *International Journal of Research in Engineering and Technology (IJRET)*, 1(4), pp. 217–220. Available at: http://journalsweb.org/siteadmin/upload/56864 IJRET014045.pdf.

Leary, M. (2012) *Introduction to Behavioral Research Methods*. Edited by 6th. Pearson.

Lee, R. M., Assante, M. J. and Conway, T. (2016) 'Analysis of the cyber attack on the Ukrainian power grid', *SANS Industrial Control Systems*, p. 23. Available at: https://ics.sans.org/media/E-ISAC_SANS_Ukraine_DUC_5.pdf.

Legg, S. and Hutter, M. (2007) 'Universal intelligence: A definition of machine intelligence', *Minds and Machines*, 17(4), pp. 391–444. doi: 10.1007/s11023-007-9079-x.

Leigh, D. and Harding, L. (2011) *Wikileaks*. London: Guardian Books.

Leopold, H. (2015) 'Cyber Situational Awareness', *Elektrotechnik und Informationstechnik*.

Leversage, D. J. and Byres, E. J. (2008) 'Estimating a system's mean time-to-compromise', *IEEE Security and Privacy*. doi: 10.1109/MSP.2008.9.

Li, Z. *et al.* (2011) 'Towards Situational Awareness of Large-scale Botnet Probing Events', *IEEE Transactions on Information Forensics and Security*, 6(1), pp. 175–188.

Lippmann, R. P. *et al.* (2016) 'Finding Malicious Cyber Discussions in Social Media', *Lincoln Laboratory Journal*, 22(1), pp. 203–209.

Liu, Y. *et al.* (2015) 'Cloudy with a Chance of Breach: Forecasting Cyber Security Incidents', *24th USENIX Security Symposium (USENIX Security 15)*, pp. 1009–1024. Available at: https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/liu.

Luetkepohl, H. (2011) 'Vector Autoregressive Models', *European University Institute*, pp. 1–22. Available at: http://ideas.repec.org/p/eui/euiwps/eco2011-30.html.

Lui, B. (2015) *Sentiment Analysis: mining sentiments, opinions, and emotions*. 1st edn. New York: Cambridge University Press.

Lukasik, S. J. (2000) 'Protecting the global information commons', *Telecommunications Policy*. Elsevier Science Ltd, 24(6), pp. 519–531.

Lukasik, S. J. (2011) 'Protecting users of the cyber commons', *Communications of the ACM*, 54(9), p. 54. doi: 10.1145/1995376.1995393.

Lütkepohl, H. (2004) 'Forecasting with VARMA Models', *European University Institute*, pp. 0–40. doi: 10.1016/S1574-0706(05)01006-2.

Mahoney, M. V. (2003a) 'A Machine Learning Approach to Detecting Attacks by Identifying Anomalies in Network Traffic', *Computer*. Available at: http://www.cs.fit.edu/~mmahoney/dist/diss.pdf,.

Mahoney, M. V. (2003b) 'Network traffic anomaly detection based on packet bytes', *Proceedings of the 2003 ACM symposium on Applied computing - SAC '03*, p. 346. doi: 10.1145/952589.952601.

Maier, H. R. *et al.* (2016) 'An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together?', *Environmental Modelling and Software*. Elsevier Ltd, 81, pp. 154–164. doi: 10.1016/j.envsoft.2016.03.014.

March, S. T. and Smith, G. F. (1995) 'Design and natural science research on information technology', *Decision Support Systems*, 15(4), pp. 251–266. doi: 10.1016/0167-9236(94)00041-2.

Marconato, G. V., Nicomette, V. and Ka??niche, M. (2012) 'Security-related vulnerability life cycle analysis', in *7th International Conference on Risks and Security of Internet and Systems, CRiSIS 2012*. doi: 10.1109/CRISIS.2012.6378954.

Massey, F. J. J. (1951) 'Kolmogorov-Smirnov Test for Goodness of Fit', *Journal of the American Statistical Association*, 46(253), pp. 68–78. doi: 10.1080/01621459.1951.10500769.

Matusitz, J. (2011) 'Social Network Theory: A Comparative Analysis of the Jewish Revolt in Antiquity and the Cyber Terrorism Incident over Kosovo', *Information Security Journal: A Global Perspective*, 20(1), pp. 34–44. doi: 10.1080/19393555.2010.544702.

Maurizio, L. (2002) 'Data integration: a theoretical perspective', *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 233–246. doi: http://doi.acm.org/10.1145/543613.543644.

Mauw, S. and Oostdijk, M. (2006) 'Foundations of attack trees', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3935 LNCS(im), pp. 186–198. doi: 10.1007/11734727_17.

McMahon, J. (2014) 'An Analysis of the Characteristics of Cyber Attacks', *Discovery*. Available at: http://computing.derby.ac.uk/ojs/index.php/da/article/download/26/43.

Mehmood, T. *et al.* (2012) 'A review of variable selection methods in Partial Least Squares

Regression', *Chemometrics and Intelligent Laboratory Systems*, pp. 62–69. doi: 10.1016/j.chemolab.2012.07.010.

MOD (2016) 'Cyber Primer', 2nd Ed.

Montgomery, D. C., Peck, E. A. and Vining, G. G. (2001) *Introduction to Linear Regression Analysis*, *Technometrics*. doi: 10.1198/tech.2007.s499.

Münz, G., Li, S. and Carle, G. (2007) 'Traffic Anomaly Detection Using K-Means Clustering', *GI/ITG Workshop MMBnet*. Available at: http://www.decom.ufop.br/menotti/rp122/sem/sem3-luciano-art.pdf.

Nam, S. Y. and Kim, H. S. (2006) 'Scanner Detection Based on Connection Attempt Success Ratio with Guaranteed False Positive and False Negative Probabilities'.

Nielsen, F. Å. (2011) 'A new ANEW: Evaluation of a word list for sentiment analysis in microblogs', in *CEUR Workshop Proceedings*, pp. 93–98.

Nigam, K., Lafferty, J. and Mccallum, A. (1999) 'Using Maximum Entropy for Text Classification', *IJCAI-99 Workshop on Machine Learning for Information Filtering*, pp. 61–67. doi: 10.1.1.63.2111.

Ning, H. *et al.* (2015) 'Cybermatics: Cyber–physical–social–thinking hyperspace based science and technology', *Future Generation Computer Systems*. doi: 10.1016/j.future.2015.07.012.

Ning, H. *et al.* (2016) 'Cybermatics: Cyber-physical-social-thinking hyperspace based science and technology', *Future Generation Computer Systems*. Elsevier B.V., 56, pp. 504–522. doi: 10.1016/j.future.2015.07.012.

Noppamas, A., Seree, C. and Kidakan, S. (2014) 'CUTOFF THRESHOLD OF VARIABLE IMPORTANCE IN PROJECTION FOR VARIABLE SELECTION', 94(3), pp. 307–322.

Norcross, J. C., Guadagnoli, E. and Prochaska, J. O. (1984) 'Factor structure of the Profile of Mood States (POMS): Two partial replications', *Journal of Clinical Psychology*, 40(5), pp. 1270–1277. doi: 10.1002/1097-4679(198409)40:5<1270::AID-JCLP2270400526>3.0.CO;2-7.

Ntoulas, a., Pzerfos, P. and Cho, J. C. J. (2005) 'Downloading textual hidden web content through keyword queries', *Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL '05)*, pp. 100–109. doi: 10.1145/1065385.1065407.

O'Connor, B. *et al.* (2010) 'From tweets to polls: Linking text sentiment to public opinion time series', *From tweets to polls: Linking text sentiment to public opinion time series*, pp. 122–129.

Olsen, P. (2013) *We Are Anonymous*. London: William Heinemann.

Olteanu, A. *et al.* (2014) 'CrisisLex: A Lexicon for Collecting and Filtering Microblogged Communications in Crises', *Proc. of the 8th International Conference on Weblogs and Social Media*, p. 376. doi: 10.1.1.452.7691.

Ozcicek, O. and Douglas McMillin, W. (1999) 'Lag length selection in vector autoregressive models: Symmetric and asymmetric lags', *Applied Economics*, 31(4), pp. 517–524. doi: 10.1080/000368499324237.

Pallant, J. (2009) 'SPSS Survival Manual'.

Pandey, P. and Snekkenes, E. A. (2014) 'Applicability of Prediction Markets in Information Security Risk Management', *2014 25th International Workshop on Database and Expert Systems Applications*, pp. 296–300. doi: 10.1109/DEXA.2014.66.

Pang, B. and Lee, L. (2006) 'Opinion Mining and Sentiment Analysis', *Foundations and Trends® in InformatioPang, B., & Lee, L. (2006). Opinion Mining and Sentiment Analysis. Foundations and Trends® in Information Retrieval, 1(2), 91–231. doi:10.1561/1500000001n Retrieval*, 1(2), pp. 91–231. doi: 10.1561/1500000001.

Pang, B. and Lee, L. (2008) 'Opinion Mining and Sentiment Analysis', *Foundations and Trends in Information Retrieval*, 2(1–2), pp. 1–135. Available at: http://www.nowpublishers.com/product.aspx?product=INR&doi=1500000001.

Panjwani, S. *et al.* (2005) 'An experimental evaluation to determine if port scans are precursors to an attack', in *Proceedings of the International Conference on Dependable Systems and Networks*. doi: 10.1109/DSN.2005.18.

Patcha, A. and Park, J. M. (2007) 'An overview of anomaly detection techniques: Existing solutions and latest technological trends', *Computer Networks*, 51(12), pp. 3448–3470. doi: 10.1016/j.comnet.2007.02.001.

Peffers, K. *et al.* (2007) 'A Design Science Research Methodology for Information Systems Research', *Journal of Management Information Systems*, 24(3), pp. 45–77. doi: 10.2753/MIS0742-1222240302.

Phillips, P. C. B. and Ouliaris, S. (1990) 'Asymptotic Properties of Residual Based Tests for Cointegration', *Econometrica*, 58(1), p. 165. doi: 10.2307/2938339.

Picard, R. W., Vyzas, E. and Healey, J. (2001) 'Toward machine emotional intelligence: Analysis of affective physiological state', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), pp. 1175–1191. doi: 10.1109/34.954607.

Plutchik, R. (1982) 'A psychoevolutionary theory of emotions', *Social Science Information*, pp. 529–553. doi: 10.1177/053901882021004003.

Pyle, D., Editor, S. and Cerra, D. D. (1999) *Data Preparation for Data Mining*, *Order A Journal On The Theory Of Ordered Sets And Its Applications*. doi: 10.1080/713827180.

Qin, X. and Lee, W. (2004) 'Attack plan recognition and prediction using causal networks', in *Proceedings - Annual Computer Security Applications Conference, ACSAC*, pp. 370–379.

Raghava, N. S., Sahgal, D. and Chandna, S. (2012) 'Classification of Botnet detection based on botnet architechture', in *Proceedings - International Conference on Communication Systems and Network Technologies, CSNT 2012*. doi: 10.1109/CSNT.2012.128.

Robert F. Engle and Kenneth F. Kroner (1995) 'Multivariate Simultaneous Generalized Arch Author (s): Robert F . Engle and Kenneth F . Kroner Published by : Cambridge University Press Stable URL : http://www.jstor.org/stable/3532933 JSTOR is a not-for-profit service that helps scholars , researchers', 11(1), pp. 122–150.

Rose, S. *et al.* (2010) 'Automatic keyword extraction', *Text Mining: Applications and Theory*, pp. 1--277. doi: 10.1002/9780470689646.ch1.

Roy, A., Kim, D. S. and Trivedi, K. S. (2010) 'Cyber security analysis using attack countermeasure trees', *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research - CSIIRW '10*, p. 1. doi: 10.1145/1852666.1852698.

Roy, S., Singh, A. K. and Sairam, A. S. (2011) 'Detecting and Defeating SQL Injection Attacks', *International Journal of Information and Electronics Engineering (IJIEE)*.

Sadoddin, R. and Ghorbani, A. (2006) 'Alert correlation survey', *Proceedings of the 2006 International Conference on Privacy, Security and Trust Bridge the Gap Between PST Technologies and Business Services - PST '06*, p. 1. doi: 10.1145/1501434.1501479.

Sanders, C. (2017) *Practical Packet Analysis: Using Wireshark to solve real-world network problems*. No Starch Press.

Sans Institute (2012) 'Shedding Light on Security Incidents Using Network'.

Sanzgiri, A., Hughes, A. and Upadhyaya, S. (2013) 'Analysis of malware propagation in Twitter', *Proceedings of the IEEE Symposium on Reliable Distributed Systems*, pp. 195–204. doi: 10.1109/SRDS.2013.28.

Schneier, B. (1999) 'Attack Trees', *SANS Network Security*, (March).

Scholz, T. and Conrad, S. (2013) 'Opinion Mining in Newspaper Articles by Entropy-Based Word Connections', *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, (October 2013), pp. 1828–1839. Available at: http://www.aclweb.org/anthology/D13-1188.

Schreck, T. and Keim, D. (2013) 'Visual Analysis of Social Media Data', *Comuputer*, 46(5), pp. 68–75.

Schwarz, G. (1978) 'Estimating the Dimension of a Model', *The Annals of Statistics*, 6(2), pp. 461–464. doi: 10.1214/aos/1176344136.

Sendi, A. S. *et al.* (2012) 'Real time intrusion prediction based on optimized alerts with Hidden Markov model', *Journal of Networks*, 7(2), pp. 311–321. doi: 10.4304/jnw.7.2.311-321.

Shah, K. and Tanvi, K. (2006) 'Disclosing Maalicious Traffic For Network Security', 7(6), pp. 1701–1706.

Shannon, C. E. (1948) 'A mathematical theory of communication', *The Bell System Technical Journal*, 27(July 1928), pp. 379–423. doi: 10.1145/584091.584093.

Shannon, C. E. (1948) 'The mathematical theory of communication.', *MD computing computers in medical practice*, pp. 306–17. doi: 10.1145/584091.584093.

Shapiro, S. and Wilk, M. (1965) 'An Analysis of Variance Test for Normality', *Biometrika*, 52(3/4), pp. 591–611. doi: 10.2307/2333709.

Sharma, A. *et al.* (2010) 'Building a social dimensional threat model from current and historic events of cyber attacks', *Proceedings - SocialCom 2010: 2nd IEEE International Conference on Social Computing, PASSAT 2010: 2nd IEEE International Conference on Privacy, Security, Risk and Trust*, pp. 981–986. doi: 10.1109/SocialCom.2010.145.

Sharma, A. *et al.* (2013) 'A social dimensional cyber threat model with formal concept analysis and

fact-proposition inference', *International Journal of Information and Computer Security*, 5(4), p. 301. doi: 10.1504/IJICS.2013.058213.

Shekhar, S. (2016) *ISIS offers hackers $10,000 a job in bid to steal government data*, *Mail Online UK*. Available at: http://www.dailymail.co.uk/indiahome/indianews/article-3416253/ISIS-offers-desi-hackers-10-000-job-Terror-group-contacts-30-000-social-media-tries-steal-government-data.html (Accessed: 10 March 2017).

Sims (1980) 'Vector Autoregression and Vector Error-Correction Models', *Chapter 5*, pp. 70–99.

Singh, N. K. and Roy, B. N. (2010) 'An Approach to Understand the End User Behavior through Log Analysis', *International Journal of Computer Applications*, 5(11), pp. 27–34. doi: 10.5120/953-1330.

Skopik, F. *et al.* (2012) 'Designing a cyber attack information system for national situational awareness', in *Communications in Computer and Information Science*, pp. 277–288.

Society, L. (2016) 'Linguistic Society of America Review Reviewed Work (s): WordNet : An Electronic Lexical Database by Christiane Fellbaum Review by : Adam Kilgarriff Published by : Linguistic Society of America Stable URL : http://www.jstor.org/stable/417141', 76(3), pp. 706–708.

Sofiyanti, N., Fitmawati, D. I. and Roza, A. A. (2015) *Stenochlaena Riauensis (Blechnaceae), A new fern species from riau, Indonesia*, *Bangladesh Journal of Plant Taxonomy*. Imperial College Press. doi: 10.1007/s13398-014-0173-7.2.

Sousan, W. L. *et al.* (2010) 'Using term extraction patterns to discover coherent relationships from open source intelligence', *Proceedings - SocialCom 2010: 2nd IEEE International Conference on Social Computing, PASSAT 2010: 2nd IEEE International Conference on Privacy, Security, Risk and Trust*, pp. 967–972. doi: 10.1109/SocialCom.2010.143.

Spearman, C. (1904) 'Proof and Disproof of Correlation', *The American Journal of Psychology*, 16(2), pp. 228–231.

Sproat, R. (2000) 'Lexical Analysis', *Handbook of Natural Language Processing*, pp. 37–57.

Stanfield, B. L. and Stanfield, L. (2016) 'Predicting Cyber Attacks : A Study of the Successes and Failures of the Intelligence Community'.

Tang, D. *et al.* (2014) 'Learning Sentiment-Specific Word Embedding', *Acl*, pp. 1555–1565.

Tartakovsky, A. G., Polunchenko, A. S. and Sokolov, G. (2013) 'Efficient Computer Network Anomaly Detection by Changepoint Detection Methods', *IEEE Journal of Selected Topics in Signal Processing*, 7(1), pp. 4–11. doi: 10.1109/JSTSP.2012.2233713.

Thesen, a. and Travis, L. E. (1990) *Introduction to Simulations*, *Needs Journal*.

Thomas, C., Sharma, V. and Balakrishnan, N. (2008) 'Usefulness of DARPA dataset for intrusion detection system evaluation', *Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security 2008*, 6973, p. 69730G. doi: 10.1117/12.777341.

Thwaites, P., Smith, J. Q. and Riccomagno, E. (2010) 'Causal analysis with Chain Event Graphs', *Artificial Intelligence*. Elsevier B.V., 174(12–13), pp. 889–909. doi: 10.1016/j.artint.2010.05.004.

TM Forum (2011) 'TM Forum Security Management Model', pp. 1–30.

Tsai, F. S. and Chan, K. L. (2007) 'Detecting Cyber Security Threats in Weblogs Using Probabilistic Models', *Lecture Notes in Computer Science 4430*, pp. 46–57. doi: 10.1007/978-3-540-71549-8_4.

Uma, M. and Padmavathi, G. (2013) 'A survey on various cyber attacks and their classification', *International Journal of Network Security*, 15(5), pp. 390–396.

United States: US Army (2010) 'Cyberspace Operations Concept Capability Plan 2016-2028', *TRADOC Pamphlet 525-7-8*, (February 2010). Available at: http://www.fas.org/irp/doddir/army/pam525-7-8.pdf.

United States Defense Force (2013) 'Joint Publication 3-12 Cyberspace Operations', *United States Defense Force*, 12(February 2013), p. 62. Available at: www.e-publishing.af.mil.

Ventzislav, I. and Lutz, K. (2005) 'A Practitioner's Guide to Lag Order Selection For VAR Impulse Response Analysis', *Studies in Nonlinear Dynamics & Econometrics*, 9(1), pp. 1–36. Available at: http://ideas.repec.org/a/bpj/sndecm/v9y2005i1n2.html.

Verleysen, M. and François, D. (2005) 'The Curse of Dimensionality in Data Mining', *Analysis*, 3512, pp. 758–770. doi: 10.1007/11494669_93.

Whiting, M. *et al.* (2015) 'VAST challenge 2014: The Kronos incident', *2014 IEEE Conference on Visual Analytics Science and Technology, VAST 2014 - Proceedings*, pp. 295–300. doi: 10.1109/VAST.2014.7042536.

Wilson, T., Wiebe, J. and Hoffman, P. (2005) 'Recognizing contextual polarity in phrase level sentiment analysis', *Acl*, 7(5), pp. 12–21. doi: 10.3115/1220575.1220619.

Wilson, T., Wiebe, J. and Hoffmann, P. (2005) 'Recognizing contextual polarity in phrase-level sentiment analysis', *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing - HLT '05*, 7(5), pp. 347–354. doi: 10.3115/1220575.1220619.

Winskel, G. (2011) 'Events, causality and symmetry', *Computer Journal*, 54(1), pp. 42–57.

Wlodarczyk, T. W. and Hacker, T. J. (2014) 'Current trends in predictive analytics of big data', *International Journal of Big Data Intelligence*. doi: http://dx.doi.org/10.1504/IJBDI.2014.066326.

Wu, Q. and Shao, Z. (2005) 'Network Anomaly Detection Using Time Series Analysis', *2008 International Conference on Telecommunications*. doi: 10.1109/ICAS-ICNS.2005.69.

Yadav, T. and Rao, A. M. (2015) 'Technical aspects of cyber kill chain', *Communications in Computer and Information Science*, 536, pp. 438–452. doi: 10.1007/978-3-319-22915-7_40.

Zetter, K. (2014) *Countdown to Zero Day*. United States: Crown Publishing Group. Available at: http://www.randomhouse.com/book/219931/countdown-to-zero-day-by-kim-zetter.

Zhang, F. and Wang, D. (2013) 'An effective feature selection approach for network intrusion detection', *Proceedings - 2013 IEEE 8th International Conference on Networking, Architecture and Storage, NAS 2013*. IEEE, pp. 307–311. doi: 10.1109/NAS.2013.49.

Zhang, Y., Jin, R. and Zhou, Z. H. (2010) 'Understanding bag-of-words model: A statistical framework', *International Journal of Machine Learning and Cybernetics*, 1(1–4), pp. 43–52. doi: 10.1007/s13042-

010-0001-0.

Zivot, E. and Wang, J. (2006) 'Vector Autoregressive Models for Multivariate Time Series', *Modeling Financial Time Series with S-PLUS®*, (1994), pp. 383–427. doi: 10.1007/978-0-387-32348-0_11.

9 APPENDICES

9.1 APPENDIX 1 : DATA PREPARATION

		Repository	Available At	Accessed
Data Acquisition	Data Acquisition on the Economic Dimension	Ikwu, R. (2018). *eneyi/MdDataFusion*. [online] GitHub	https://github.com/eneyi/MdDataFusion/tree/master/DataAquisition/EconomicDimension	Accessed 10 Feb. 2019
	Data Acquisition on the Physical Dimension		https://github.com/eneyi/MdDataFusion/tree/master/DataAquisition/PhysicalDimension	
	Data Acquisition on the Social Dimension		https://github.com/eneyi/MdDataFusion/tree/master/DataAquisition/SociallDimension	
Data Transformation	Data Transformation on the Economic Dimension	Ikwu, R. (2018). *eneyi/MdDataFusion*. [online] GitHub	https://github.com/eneyi/MdDataFusion/tree/master/DataTransformation/EconomicDimension	Accessed 10 Feb. 2019
	Data Transformation on the Physical Dimension		https://github.com/eneyi/MdDataFusion/tree/master/ DataTransformation /PhysicalDimension	
	Data Transformation on the Social Dimension		https://github.com/eneyi/MdDataFusion/tree/master/ DataTransformation /SociallDimension	