Hindawi
Security and Communication Networks
Volume 2018, Article ID 8028960, 10 pages
https://doi.org/10.1155/2018/8028960

Research Article

An Effective Classification Approach for Big Data Security Based on GMPLS/MPLS Networks

Sahel Alouneh,[1] Feras Al-Hawari,[1] Ismail Hababeh ,[1] and Gheorghita Ghinea [2]

[1] *German Jordanian University, Jordan*
[2] *Brunel University, UK*

Correspondence should be addressed to Ismail Hababeh; ismail.hababeh@gju.edu.jo

Received 5 February 2018; Accepted 9 May 2018; Published 13 June 2018

Academic Editor: Petros Nicopolitidis

The need for effective approaches to handle big data that is characterized by its large volume, different types, and high velocity is vital and hence has recently attracted the attention of several research groups. This is especially the case when traditional data processing techniques and capabilities proved to be insufficient in that regard. Another aspect that is equally important while processing big data is its security, as emphasized in this paper. Accordingly, we propose to process big data in two different tiers. The first tier classifies the data based on its structure and on whether security is required or not. In contrast, the second tier analyzes and processes the data based on volume, variety, and velocity factors. Simulation results demonstrated that using classification feedback from a MPLS/GMPLS core network proved to be key in reducing the data evaluation and processing time.

1. Introduction

Big data is a new term that refers not only to data of big size, but also to data with unstructured characteristic types (i.e., video, audio, unstructured text, and social media information). The demand for solutions to handle big data issues has started recently by many governments' initiatives, especially by the US administration in 2012 when it announced the big data research and development initiative [1]. The initiative aims at exploring proper and efficient ways to use big data in solving problems and threats facing the nation, government, and enterprise. It is also worth noting that analyzing big data information can help in various fields such as healthcare, education, finance, and national security.

Potential challenges for big data handling consist of the following elements [3]:

(i) **Analysis**: this process focuses on capturing, inspecting, and modeling of data in order to extract useful information.

(ii) **Treatment and conversion**: this process is used for the management and integration of data collected from different sources to achieve useful presentation, maintenance, and reuse of data.

(iii) **Searching**: this process is considered the most important challenge in big data processing as it focuses on the most efficient ways to search inside data that it is big and not structured on one hand and on the timing and correctness of the extracted searched data on the other hand.

(iv) **Storage**: this process includes best techniques and approaches for big data organization, representation, and compression, as well as the hierarchy of storage and performance.

(v) **Visualization**: this process involves abstracting big data and hence it helps in communicating data clearly and efficiently.

(vi) **Security and sharing**: this process focuses on data privacy and encryption, as well as real-time analysis of coded data, in addition to practical and secure methods for data sharing.

The increasing trend of using information resources and the advances of data processing tools lead to extend usage of big data. The extensive uses of big data bring different challenges, among them are data analysis, treatment and conversion, searching, storage, visualization, security, and privacy.

Big data security and privacy are potential challenges in cloud computing environment as the growing usage of big data leads to new data threats, particularly when dealing with sensitive and critical data such as trade secrets, personal and financial information. Any loss that could happen to this data may negatively affect the organization's confidence and might damage their reputation. Moreover, moving big data within different clouds that have different levels of sensitivity might expose important data to threats.

However, the traditional methods do not comply with big data security requirements where tremendous data sets are used. Consequently, new big data security and privacy techniques are required to overcome data threats and its risk management.

Therefore, this research aims at exploring and investigating big data security and privacy threats and proposes twofold approach for big data classification and security to minimize data threats and implements security controls during data exchange. The primary contributions of this research for the big data security and privacy are summarized as follows:

(i) Classifying big data according to its structure that help in reducing the time of applying data security processes.

(ii) Using of data-carrying technique, Multiprotocol Label Switching (MPLS) to achieve high-performance telecommunication networks.

(iii) Transferring big data from one node to another based on short path labels rather than long network addresses to avoid complex lookups in a routing table.

(iv) Using labels in order to differentiate between traffic information that comes from different networks.

(v) Analyzing and processing big data at Networks Gateways that help in load distribution of big data traffic and improve the performance of big data analysis and processing procedures.

The rest of the paper is organized as follows. In Section 2, the related work that has been carried out on big data in general with a focus on security is presented. In Section 3, the proposed approach for big data security using classification and analysis is introduced. In Section 4, the validation results for the proposed method are shown. Finally, in Section 5, conclusions and future work are provided.

2. Related Work

Research work in the field of big data started recently (in the year of 2012) when the White House introduced the big data initiative [1]. The research on big data has so far focused on the enhancement of data handling and performance. On the other hand, handling the security of big data is still evolving and just started to attract the attention of several research groups. In this section, we present and focus on the main big data security related research work that has been proposed so far. Indeed, our work is different from others in considering the network core as a part of the big data classification process. Furthermore and to the best of our knowledge, the proposed approach is the first to consider the use of a Multiprotocol Label Switching (MPLS) network and its characteristics in addressing big data QoS and security.

Authors in [2] propose an attribute selection technique that protects important big data. The technique analyzes big data by extracting valuable content that needs protection. It mainly extracts information based on the relevance factor. However, it does not support or tackle the issue of data classification; i.e., it does not discuss handling different data types such as images, regular documents, tables, and real-time information (e.g., VoIP communications). In [3], the authors investigated the security issues encountered by big data when used in cloud networks. The main issues covered by this work are network security, information security, and privacy. The authors in [4] developed a new security model for accessing distributed big data content within cloud networks. The proposed security framework focuses on securing autonomous data content and is developed in the G-Hadoop distributed computing environment.

The challenge to legitimately use big data while considering and respecting customer privacy was interestingly studied in [5]. In related work [6], its authors considered the security awareness of big data in the context of cloud networks with a focus on distributed cloud storages via STorage-as-a-Service (STaaS). In [7], they also addressed big data issues in cloud systems and Internet of Things (IoT). Specifically, they summarized and analyzed the main results obtained when external integrity verification techniques are used for big data security within a cloud environment. In [8], they proposed to handle big data security in two parts. The first part challenges the credibility of security professionals' discourses in light of the knowledge that they apparently mobilize, while the second part suggests a series of conceptual interchanges around data, relationships, and procedures to address some of the restrictions of current activities with the big data security assemblage.

Furthermore, in [9], they considered the security of real-time big data in cloud systems. The work is based on a multilayered security paradigm that can protect data in real time at the following security layers: firewall and access control, identity management, intrusion prevention, and convergent encryption. Another work that targets real-time content is presented in [10], in which a semantic-based video organizing platform is proposed to search videos in big data volumes. The proposed technique uses a semantic relational network model to mine and organize video resources based on their associations, while the authors in [11] proposed a Dynamic Key Length based Security Framework (DLSeF) founded on a common key resulting from synchronized prime numbers. The key is dynamically updated in short intervals to prevent man in the middle attacks. In contrast, the authors in [12] focused on the big data multimedia content problem within a cloud system. They proposed a novel approach using Semantic-Based Access Control (SBAC) techniques for acquiring secure financial services.

Moreover, the work in [13] focused on the privacy problem and proposed a data encryption method called Dynamic

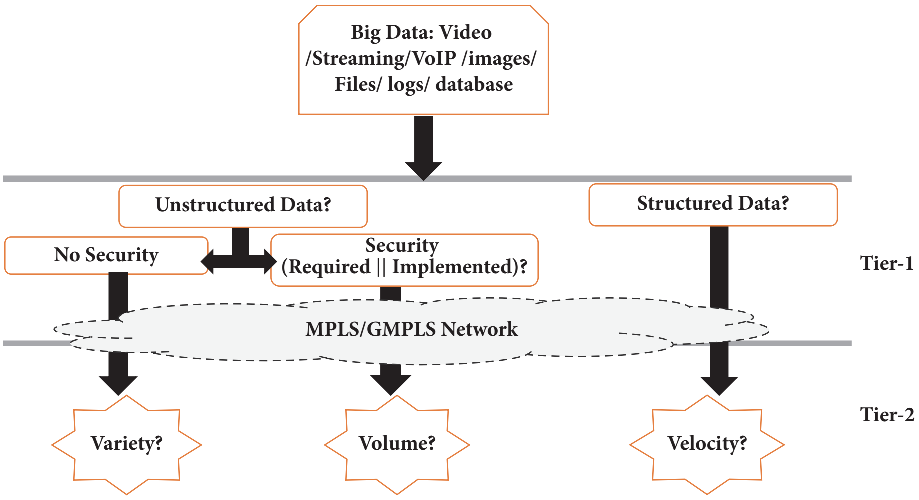


Figure 1: A flow chart of the general architecture for our approach.

Data Encryption Strategy (D2ES). The method selectively encodes information using privacy classification methods under timing constraints. Besides that, other research studies [14–24] have also considered big data security aspects and solutions.

3. Classification Approach

Big data can contain different kinds of information such as text, video, financial data, and logs, as well as secure or insecure information. Thus, the treatment of these different sources of information should not be the same. Furthermore, the proposed classification method should take the following factors into consideration [5].

Velocity: the speed of data generation and processing.

Volume: the size of data generated and storage space required.

Variety: the category of data and its characteristics.

The proposed classification algorithm is concerned with processing secure big data. Since handling secure data is different than plaintext data, the following factors should be taken into consideration in our algorithm.

Confidentiality: the confidentiality factor is related to whether the data should be encrypted or not. Therefore, with security in mind, big data handling for encrypted content is not a simple task and thus requires different treatment.

Authentication: some big data may require authentication, i.e., protection of data against modification. In addition, authentication deals with user authentication and a Certification Authority (CA).

Other security factors such as Denial of Service (DoS) protection and Access Control List (ACL) usage will also be considered in the proposed algorithm.

In the following subsections, the details of the proposed approach to handle big data security are discussed.

3.1. General Architecture of the Classification Approach. In the proposed approach, big data is processed by two hierarchy tiers. The role of the first tier (Tier 1) is concerned with the classification of the big data to be processed. In other words, this tier decides first on whether the incoming big data traffic is structured or unstructured. Then, it checks the type of security service that is applied on the data, i.e., whether encryption is applied or not on the processed data, or if authentication is implemented or required on the processed data. The second tier (Tier 2) decides on the proper treatment of big data based on the results obtained from the first tier, as well as based on the analysis of velocity, volume, and variety factors. A flow chart for the general architecture of the proposed method is shown in Figure 1.

3.2. Tier 1: Data Classification. In this subsection, the algorithm used to classify big data information (Tier 1) (i.e., whether data is structured or unstructured and whether security is applied or not) is presented.

Before processing the big data, there should be an efficient mechanism to classify it on whether it is structured or not and then evaluate the security status of each category. The proposed algorithm relies on different factors for the analysis and is summarized as follows:

(i) Data Source and Destination (DSD): data source as well as destination may initially help to guess the structure type of the incoming data. This factor is used as a prescanning stage in this algorithm, but it is not a decisive factor. Hence, it helps to accelerate data classification without the need to perform a detailed analysis of incoming data. However, the algorithm uses a controlling feedback for updating.

(ii) Data Header information (DH): it has been assumed that incoming data is encapsulated in headers. Therefore, header information can play a significant role in data classification. For example, the IP networking traffic header contains a Type of Service (ToS) field, which gives a hint on the type of data (real-time data, video-audio data, file data, etc.). In addition,

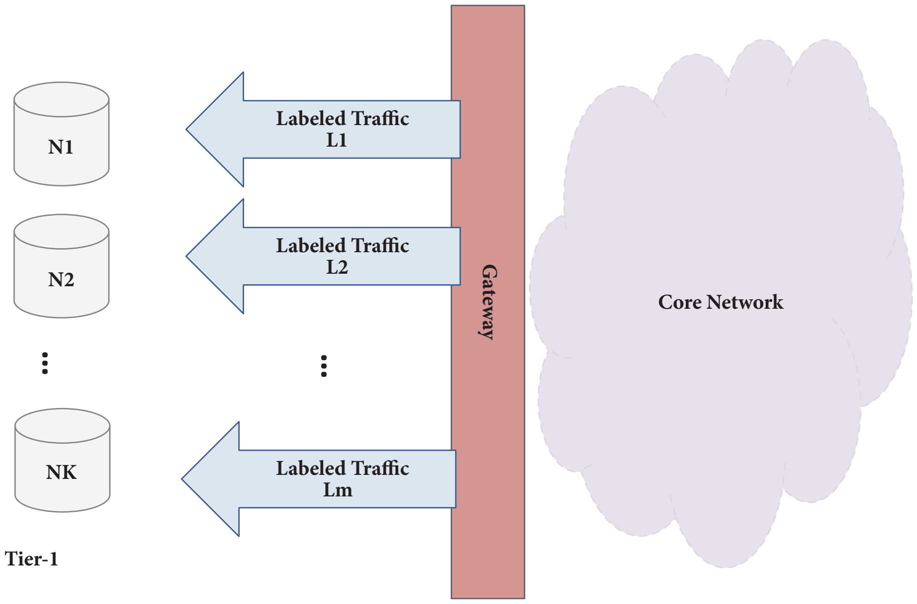


Figure 2: Tier 1 structure.

the *protocol* field indicates the upper layers, e.g., UDP, TCP, ESP security, AH security, etc.

Furthermore, the Tier 1 classification process can be enhanced by using traffic labeling. In other words, Labels (L) can be used to differentiate or classify incoming traffic data. Therefore, we assume that the network infrastructure core supports Multiprotocol Label Switching (MPLS) or the Generalized Multiprotocol Label Switching (GMPLS) [25], and thus labels can be easily implemented and mapped. Traffic that comes from different networks is classified at the gateway of the network responsible to analyze and process big data. An MPLS network core uses labels to differentiate traffic information. The MPLS header is four bytes long and the labels are created from network packet header information. The labels can carry information about the type of traffic (i.e., real time, audio, video, etc.). It is worth noting that label(s) is built from information available at (DH) and (DSD). In the Tier 1 structure shown in Figure 2, the gateway is responsible for categorizing the incoming traffic into labels called *labeled traffic* (L_m). At this stage, the traffic structure (i.e., structured or unstructured) and type (i.e., security services applied or required, or no security) should be identified.

Consequently, the gateway is responsible for distributing the labeled traffic to the appropriate node (N_K) for further analysis and processing at Tier 2. If the traffic has no security requirements, or not required, the gateway should forward that traffic to the appropriate node(s) that is/are designated to process traffic (i.e., some nodes are responsible to process traffic with requirements for security services, and other nodes are designated to process traffic data with no security requirements). This approach as will be shown later on in this paper helps in load distribution for big data traffic, and hence it improves the performance of the analysis and processing steps.

3.3. Tier 2: Data Classification. At this stage, Tier 2 takes care of the analysis and processing of the incoming labeled big data traffic which has already been screened by Tier 1. Moreover, Tier 2 is responsible for evaluating the incoming traffic according to the Velocity, Volume, and Variety factors. Each node is also responsible for analyzing and processing its assigned big data traffic according to these factors.

Next, the node internal architecture and the proposed algorithm to process and analyze the big data traffic are presented.

3.3.1. Node Architecture. The main components of Tier 2 are the nodes (i.e., $N_1, N_2, \ldots, N_K$). The internal node architecture of each node is shown in Figure 3. An internal node consists of a *Name_Node* and *Data_Node(s)*, while the incoming labeled traffic is processed and analyzed for security services based on three factors: Volume, Velocity, and Variety.

Each Tier 2 node applies Algorithms 1 and 2 when processing big data traffic. The first algorithm (Algorithm 1) decides on the security analysis and processing based on the Volume factor, whereas the second algorithm (Algorithm 2) is concerned with Velocity and Variety factors.

Algorithms 1 and 2 can be summarized as follows:

(i) The two-tier approach is used to filter incoming data in two stages before any further analysis.

(ii) Tier 1 is responsible to filter incoming data by deciding on whether it is structured or nonstructured. Thus, security analysis will be more likely to be applied on structured data or otherwise based on selection.

(iii) Tier 2 is responsible to process and analyze big data traffic based on Volume, Velocity, and Variety factors. The core idea in the proposed algorithms depends on the use of labels to filter and categorize the processed

```
Step 1: Receive (Labeled Big Data Traffic_, Gateway_number, Factor_Vol,
Security_Servic_);
          Function for getting Big Data traffic by Name_node
          Where:
          Labeled Big Data Traffic_ → traffic label (TL), Data (D),
          Gateway_number →Value of Gateway sending labeled Big Data traffic (GN)
          Factor_Vol → Volume factor type (V):
                    Assumptions:
                              (i) Real time data is assigned different label than file transfer data and
                                 thus the label value should indicate the Volume size
                              (ii) Real time data are usually assumed less than 150 bytes per packet.
          Security_Service_ → Confidentiality (C), Authentication (Auth) are examined by
the label value of each traffic.
Step 2: Forward (Data_node_, Security_Service)
          Function for distributing the labeled traffic for the designated data_node(s) with
security service assignment:
          Security_service required or applied?
          True then go to data_node(s) assigned to analyze and process security services.
Step 3: Data_node_Checking (packet_headers, labels)
          Performs header and label information checking:
          → Label_checking performs label examination
          → Header_checking performs header examination:
                    Assumptions: secured data comes with extra header size such as ESP header
Step 4: Go to step 1
```

Algorithm 1: Big data security analysis and processing based on volume.

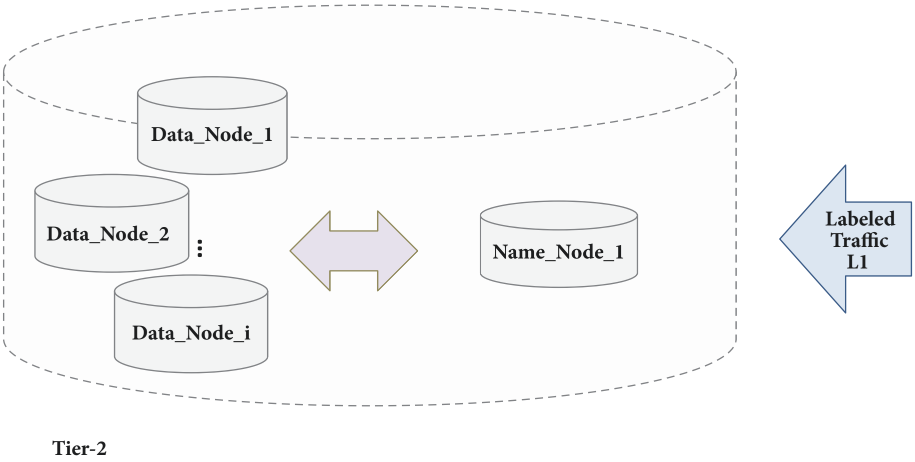


Figure 3: Node architecture.

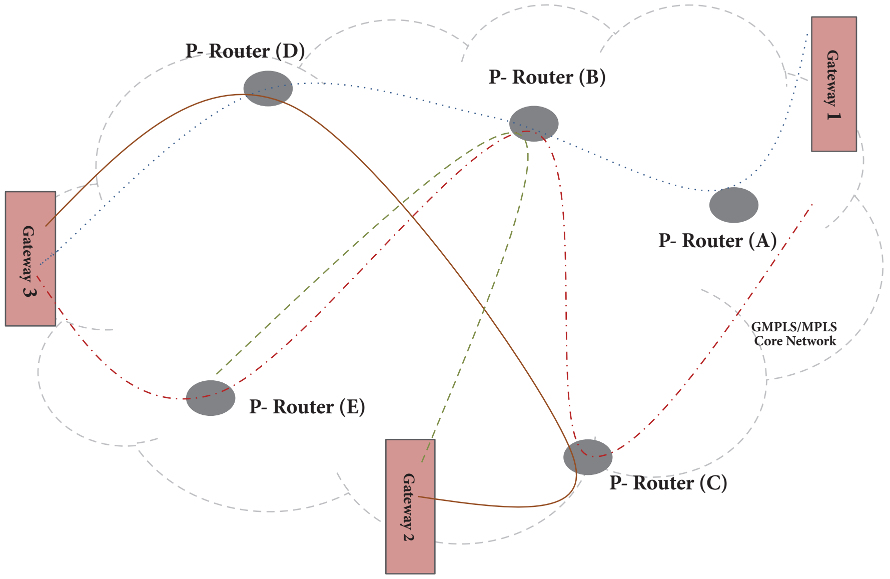
big data traffic. The network core labels are used to help tier node(s) to decide on the type and category of processed data. Thus, the use of MPLS labels reduces the burden on tier node(s) to do the classification task and therefore this approach improves the performance. On the other hand, if nodes do not support MPLS capabilities, then classification with regular network routing protocols will consume more time and extra bandwidth.

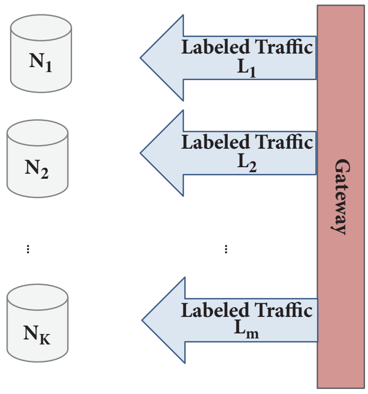
3.3.2. Network Labeling and Mapping with Big Data Node Architecture. So far, the node architecture that is used for processing and classifying big data information is presented. However, the proposed approach also requires feedback from the network in order to classify the processed data. Our assumption here is the availability of an underlying network core that supports data labeling. As mentioned in previous section, MPLS is our preferred choice as it has now been adopted by most Internet Service Providers (ISPs). The MPLS header and labeling distribution protocols make the classification of big data at processing node(s) more efficient with regard to performance, design, and implementation.

Figure 4 illustrates the mapping between the network core, which is assumed here to be a Generalized Multiprotocol Label Switching (GMPLS) or MPLS network. The GMPLS extends the architecture of MPLS by supporting switching

```
Step 1: Receive (Labeled Big Data Traffic_, Gateway_number, Factor_Vel_Var,
DSD_prob);
        Function for getting Big Data traffic by Name_node
        Where:
        Labeled Big Data Traffic_ → traffic label (TL), Data (D),
        Gateway_number →Value of Gateway sending labeled Big Data traffic (GN)
        Factor_Value → Factor type ( Var || Vel):
                Assumptions:
                     (i) Data Source and Destination (DSD) information are used and
                        integrated in Labeled traffic.
                     (ii) Data source indicates the type of data (e.g., streaming data,
                        structured nor not, ect.).
                     (iii) DSD_prob is the probability of the Velocity or Variety data
                        type.
Step 2: Forward (Data_node_, Security_Service)
        Function for distributing the labeled traffic for the designated data node(s) with
security service assignment:
        Based on the DSD probability value(s), decision is made on the security service?
        True then go to data_node(s) assigned to analyze and process security services.
Step 3: Go to step 1
```

Algorithm 2: Big data security analysis and processing based on velocity and variety.





Outgoing Label	QoS/Big Data Classification	Interface Out	Interface In	Incoming Label
L1:30	VoIP/Secure/Previous node: Gateway 1/Destination: Gateway 3	X3	Y3	25
L2:40	Streaming/Non-Secure/Previous node: P-router (C)/Destination: Gateway 3	X2	Y1	12
. . .	. . . .	. . .	. . .	. . .
Lm:55	File-Transfer/Secure/Previous node: P-router (E)/Destination: Gateway 2	X0	Y1	19

Figure 4: Network core/big data node mapping.

for wavelength, space, and time switching in addition to the packet switching. The core network consists of provider routers called here P routers and numbered A, B, etc. The GMPLS/MPLS network is terminated by complex provider Edge routers called here in this work Gateways. The Gateways are responsible for completing and handling the mapping in between the node(s), which are responsible for processing the big data traffic arriving from the core network. In addition, the gateways outgoing labeled traffic is the main factor used for data classification that is used by Tier 1 and Tier 2 layers. Algorithms 1 and 2 are the main pillars used to perform the mapping between the network core and the big data processing nodes.

3.3.3. Security Analysis. The proposed architecture supports security features that are inherited from the GMPLS/MPLS architecture, which are presented below:

Traffic Separation. The use of the GMPLS/MPLS core network provides traffic separation by using Virtual Private Network (VPN) labeling and the stacking bit (S) field that is supported by the GMPLS/MPLS headers. To illustrate more, traffic separation is an essential needed security feature. For example, if two competing companies are using the same ISP, then it is very crucial not to mix and forward the traffic between the competing parties. Nevertheless, traffic separation can be achieved by applying security encryption techniques, but this will clearly affect the performance of the network due to the overhead impact of extra processing and delay. In the proposed GMPLS/MPLS implementation, this overhead does not apply because traffic separation is achieved automatically by the use of MPLS VPN capability, and therefore our solution performs better in this regard.

Hiding Network Interior Design and Structure. One basic feature of GMPLS/MPLS network design and structure is that the incoming or outgoing traffic does not require the knowledge of participating routers inside the core network. Actually, the traffic is forwarded/switched internally using the labels only (i.e., not using IP header information). Therefore, attacks such as IP spoofing and Denial of Service (DoS) can efficiently be prevented.

Reliability and Availability. Having reliable data transfer, availability, and fast recovery from failures are considered important protection requirements and thus improve the security. Using an underlying network core based on a GMPLS/MPLS architecture makes recovery from node or link failures fast and efficient. Many recovery techniques in the literature have shown that reliability and availability can greatly be improved using GMPLS/MPLS core networks [26].

Big Data Encryption and Authentication. GMPLS/MPLS are not intended to support encryption and authentication techniques as this can downgrade the performance of the network. Therefore, security implementation on big data information is applied at network edges (e.g., network gateways and the big data processing nodes). However, Virtual Private Networks (VPNs) capabilities can be supported because of the use of GMPLS/MPLS infrastructure. The VPN capability that can be supported in this case is the traffic separation, but with no encryption. In case encryption is needed, it will be supported at nodes using appropriate encryption techniques.

4. Evaluations and Results

The main improvement of our proposed work is the use of high speed networking protocol (i.e., GMPLS/MPLS) as an underlying infrastructure that can be used by processing node(s) at network edges to classify big data traffic. Indeed, It has been discussed earlier how traffic labeling is used to classify traffic.

Now, our goal in this section is to test by simulations and analyze the impact of using the labeling approach on improving the classification of big data and thus improving the security. Therefore, in this section, simulation experiments have been made to evaluate the effect of labeling on performance. The performance factors considered in the simulations are bandwidth overhead, processing time, and data classification detection success. The simulations were conducted using the NS2 simulation tool (NS-2.35).

We have chosen different network topologies with variable distances between nodes ranging from 100m to 4000Km in the context of wired networks (LAN, WAN, MAN). Indeed, the purpose of making the distance between nodes variable is to help measuring the distance effect on processing time. The employed protocol as a routing agent for routing is the Open Shortest Path First (OSPF), while the simulation takes into consideration different scenarios for traffic rate and variable packets sizes, as detailed in Table 1.

Figure 5 shows the effect of labeling on the network overhead. The network overhead is here defined as the overhead needed to communicate big data traffic packets through the network core until being processed by edge node(s). Communication parameters include traffic engineering-explicit routing for reliability and recovery, traffic engineering- for traffic separation VPN, IP spoofing. It can be clearly noticed the positive impact of using labeling in reducing the network overhead ratio. The type of traffic analyzed in this simulation is files logs, and the simulated data size ranges from a traffic size of 100 Mbytes to 2000 Mbytes.

The effect of labeling implementation on the total nodal processing time for big data analysis has been shown in Figure 6. It can be noticed that the total processing time has been reduced significantly. Moreover, it also can be noticed that processing time increases as the traffic size increases; however, the increase ratio is much lower in the case of labeling compared to that with no labeling. The type of traffic used in the simulation is files logs. In addition, the simulated network data size ranges from 100 M bytes to 2000 M bytes.

In Figure 7, total processing time simulation has been measured again but this time for a fixed data size (i.e., 500 M bytes) and a variable data rate that ranges from 10 Mbps to 100 Mbps. The type of traffic used in the simulation is files

Table 1: Simulation parameters.

Parameter	Type/Value
Traffic Type	Long files, documents VoIP traffic
Simulation area: Distance between nodes	100m, 1 Km 10Km, 100km 1000Km, 2000Km
Routing Agents	OSPF
Sending Frequency	100 Pkts/s 300 Pkts/s 600 Pkts/s 1000 Pkts/s
Packet size	100 bytes, 200 bytes 500 bytes, 1 Kbytes
Network Infrastructure type	Wired, LAN, WAN, MAN
Number of nodes	6, 10, 6, 15, 20
Traffic	TCP, UDP
Traffic throughput	CBR
Simulation time	500s

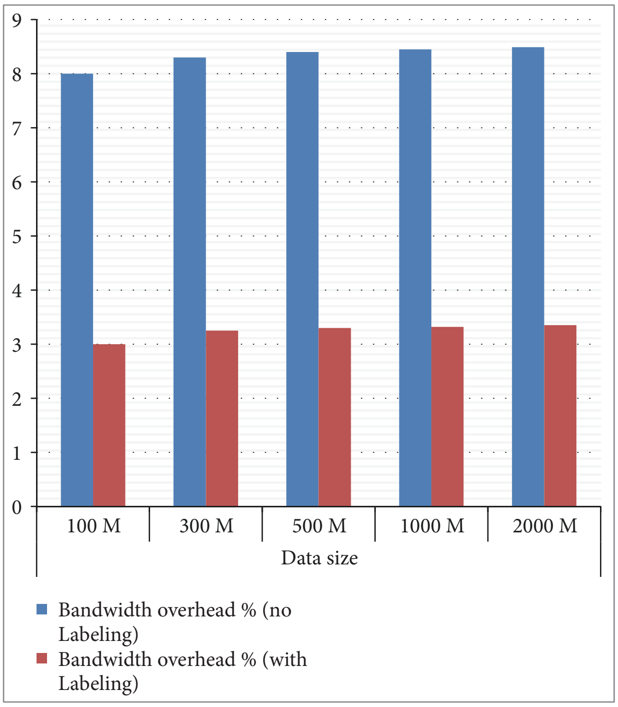


Figure 5: The ratio effect of labeling use on network overhead.

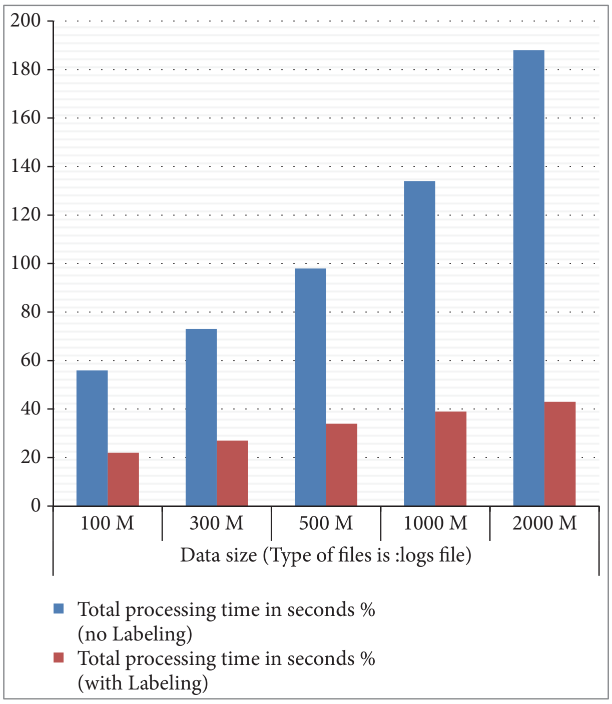


Figure 6: Total processing time in seconds for variable big data size.

logs. As can be noticed from the obtained results, the labeling methodology has lowered significantly the total processing time of big data traffic. Moreover, it also can be noticed the data rate variation on the total processing with labeling is very little and almost negligible, while without labeling the variation in processing time is significant and thus affected by the data rate increase.

We also have conducted a simulation to measure the big data classification using the proposed labeling method and compare it with the regular method when no labeling is used as shown in Figure 8. The type of data used in the simulation is VoIP, documents, and images. It can be clearly seen that the proposed method lowers significantly the processing time for data classification and detection. We also simulated in Figure 9 the effectiveness of our method in detecting IP spoofing attacks for variable packet sizes that range from 80 bytes (e.g., for VoIP packets) to 1000 bytes (e.g., for documents packet types). Our proposed method has more success time compared to those when no labeling is used.

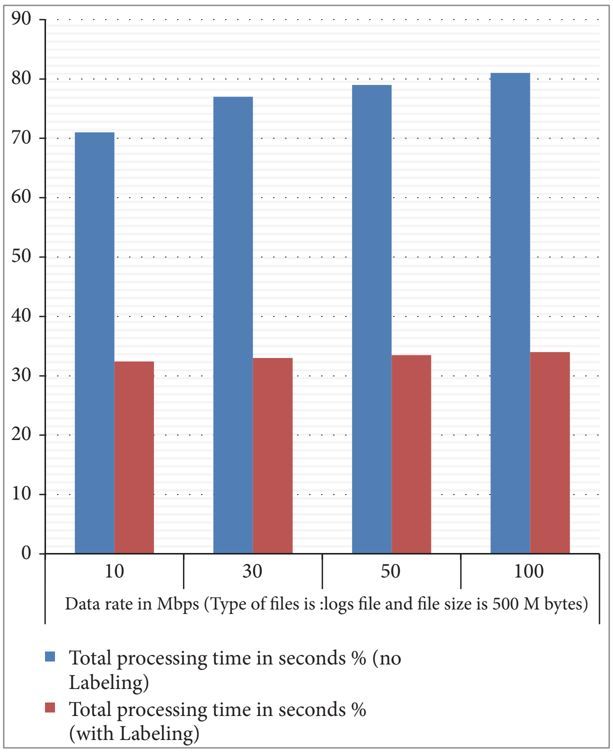


FIGURE 7: Total processing time in seconds for variable network data rate.

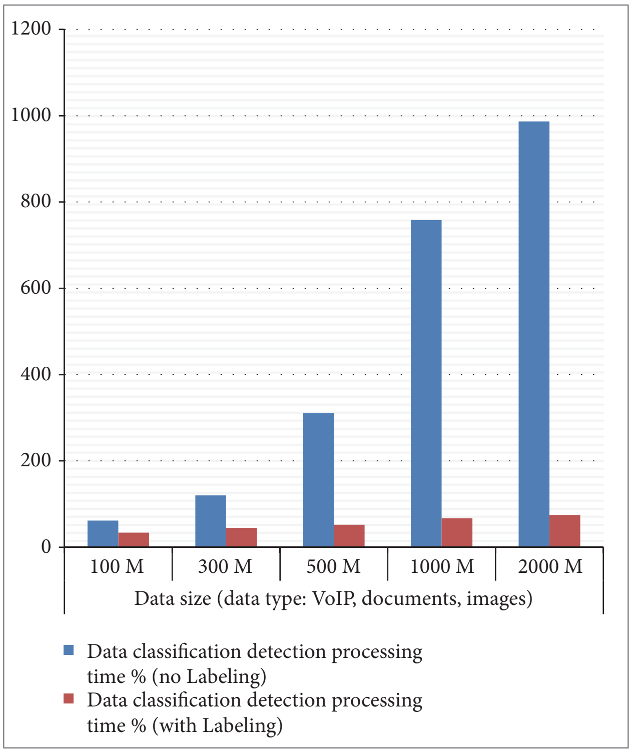


FIGURE 9: Data classification processing time in seconds for variable data types.

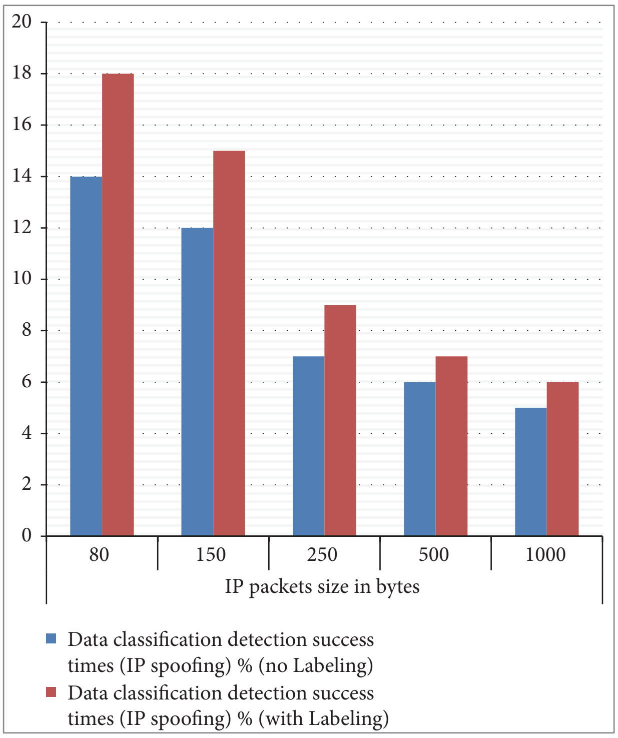


FIGURE 8: Data classification detection success time of IP spoofing attacks.

5. Conclusion and Future Work

In this paper, a new security handling approach was proposed for big data. The proposed method is based on classifying big data into two tiers (i.e., Tier 1 and Tier 2). The classification requires a network infrastructure that supports GMPLS/MPLS capabilities. The GMPLS/MPLS simplifies the classification by providing labeling assignments for the processed big data traffic. The obtained results show the performance improvements of the classification while evaluating parameters such as detection, processing time, and overhead. Future work on the proposed approach will handle the visualization of big data information in order to provide abstract analysis of classification. Furthermore, more security analysis parameters are to be investigated such as integrity and real time analysis of big data.

Data Availability

All-Schemes.TCL and Labeling-Tier.c files should be incorporated along with other MPLS library files available in NS2 and then run them for the intended parameters to generated simulation data. Data can be accessed at https://data.mendeley.com/datasets/7wkxzmdpft/2.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Executive Office of the President, “Big Data Across the Federal Government,” WH official website, March 2012.

[2] I. Narasimha, A. Sailaja, and S. Ravuri, “Security Issues Associated with Big Data in Cloud Computing,” *International Journal of Network Security and Its Applications*, vol. 6, no. 3, pp. 45–56, 2014.

[3] S.-H. Kim, N.-U. Kim, and T.-M. Chung, “Attribute relationship evaluation methodology for big data security,” in *Proceedings of the 2013 3rd International Conference on IT Convergence and Security, ICITCS 2013*, China, December 2013.

[4] J. Zhao, L. Wang, J. Tao et al., “A security framework in G-Hadoop for big data computing across distributed cloud data centres,” *Journal of Computer and System Sciences*, vol. 80, no. 5, pp. 994–1007, 2014.

[5] G. Lafuente, “The big data security challenge,” *Network Security*, vol. 2015, no. 1, pp. 12–14, 2015.

[6] K. Gai, M. Qiu, and H. Zhao, “Security-Aware Efficient Mass Distributed Storage Approach for Cloud Systems in Big Data,” in *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity)*, pp. 140–145, New York, NY, Usa, April 2016.

[7] C. Liu, C. Yang, X. Zhang, and J. Chen, “External integrity verification for outsourced big data in cloud and IoT: a big picture,” *Future Generation Computer Systems*, vol. 49, pp. 58–67, 2015.

[8] A. Claudia and T. Blanke, “The (Big) Data-security assemblage: Knowledge and critique,” *Big Data Security*, vol. 2, p. 12, July 2015.

[9] V. Chang and M. Ramachandran, “Towards Achieving Data Security with the Cloud Computing Adoption Framework,” *IEEE Transactions on Services Computing*, vol. 9, no. 1, pp. 138–151, 2016.

[10] Z. Xu, Y. Liu, L. Mei, C. Hu, and L. Chen, “Semantic based representing and organizing surveillance big data using video structural description technology,” *The Journal of Systems and Software*, vol. 102, pp. 217–225, 2015.

[11] D. Puthal, S. Nepal, R. Ranjan, and J. Chen, “A Dynamic Key Length Based Approach for Real-Time Security Verification of Big Sensing Data Stream,” in *WISE 2015: 16th International Conference*, pp. 93–108, Miami, FL, USA, November 1-3, 2015.

[12] Y. Li, K. Gai, Z. Ming, H. Zhao, and M. Qiu, “Intercrossed access controls for secure financial services on multimedia big data in cloud systems,” *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 12, no. 4s, article no. 67, 2016.

[13] K. Gai, M. Qiu, H. Zhao, and J. Xiong, “Privacy-Aware Adaptive Data Encryption Strategy of Big Data in Cloud Computing,” in *Proceedings of the 3rd IEEE International Conference on Cyber Security*, pp. 273–278, China, June 2016.

[14] V. Chang, Y.-H. Kuo, and M. Ramachandran, “Cloud computing adoption framework: A security framework for business clouds,” *Future Generation Computer Systems*, vol. 57, pp. 24–41, 2016.

[15] H. Liang and K. Gai, “Internet-Based Anti-Counterfeiting Pattern with Using Big Data in China,” *The IEEE International Symposium on Big Data Security on Cloud*, pp. 1387–1392, 2015.

[16] Z. Yan, W. Ding, X. Yu, H. Zhu, and R. H. Deng, “Deduplication on Encrypted Big Data in Cloud,” in *IEEE Transactions on Big Data*, vol. 2, pp. 138–150, 2016.

[17] A. Gholami and E. Laure, “Big Data Security and Privacy Issues in the Coud,” *International Journal of Network Security and Its Applications (IJNSA)*, vol. 8, no. 1, 2016.

[18] Y. Li, K. Gai, L. Qiu, M. Qiu, and H. Zhao, “Intelligent cryptography approach for secure distributed big data storage in cloud computing,” *Information Sciences*, vol. 387, pp. 103–115, 2017.

[19] A. Narayanan, J. Huey, and E. W. Felten, “A Precautionary Approach to Big Data Privacy,” in *Data Protection on the Move*, vol. 24 of *Law, Governance and Technology Series*, pp. 357–385, Springer Netherlands, Dordrecht, 2016.

[20] S. Kang, B. Veeravalli, and K. M. M. Aung, “A Security-Aware Data Placement Mechanism for Big Data Cloud Storage Systems,” in *Proceedings of the 2nd IEEE International Conference on Big Data Security on Cloud*, pp. 327–332, New York, NY, USA, April 2016.

[21] J. Domingo-Ferrer and J. Soria-Comas, “Anonymization in the Time of Big Data,” in *Privacy in Statistical Databases*, vol. 9867 of *Lecture Notes in Computer Science*, pp. 57–68, Springer International Publishing, 2016.

[22] Y.-S. Jeong and S.-S. Shin, “An efficient authentication scheme to protect user privacy in seamless big data services,” *Wireless Personal Communications*, vol. 86, no. 1, pp. 7–19, 2016.

[23] R. F. Babiceanu and R. Seker, “Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook,” *Computers in Industry*, vol. 81, pp. 128–137, 2016.

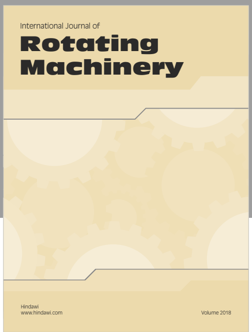
[24] Z. Xu, Z. Wu, Z. Li et al., “High Fidelity Data Reduction for Big Data Security Dependency Analyses,” in *the 2016 ACM SIGSAC Conference*, pp. 504–516, New York, NY, USA, October 2016.

[25] S. Alouneh, S. Abed, M. Kharbutli, and B. J. Mohd, “MPLS technology in wireless networks,” *Wireless Networks*, vol. 20, no. 5, pp. 1037–1051, 2014.

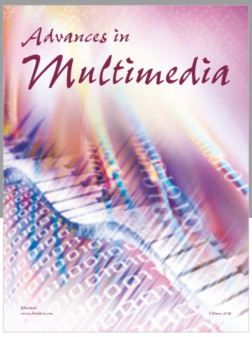
[26] S. Alouneh, A. Agarwal, and A. En-Nouaary, “A novel path protection scheme for MPLS networks using multi-path routing,” *Computer Networks*, vol. 53, no. 9, pp. 1530–1545, 2009.

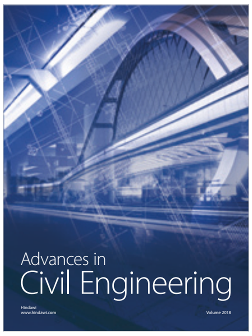
Journal of
Engineering

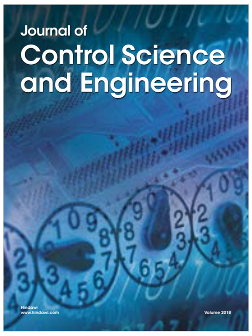
The Scientific
World Journal


International Journal of
Rotating
Machinery

Journal of
Sensors

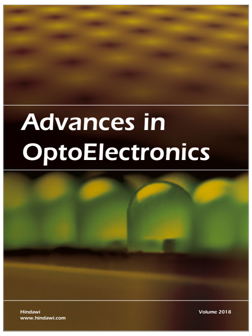

Advances in
Multimedia


Advances in
Civil Engineering


Journal of
Control Science
and Engineering

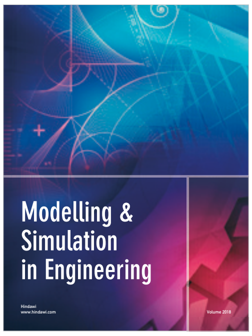
Journal of
Robotics

Journal of
Electrical and Computer
Engineering

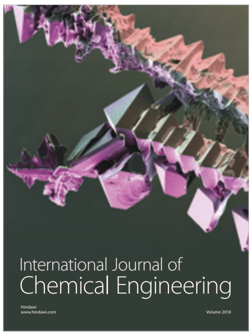

Advances in
OptoElectronics

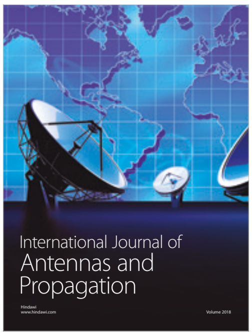
VLSI Design

International Journal of
Navigation and
Observation

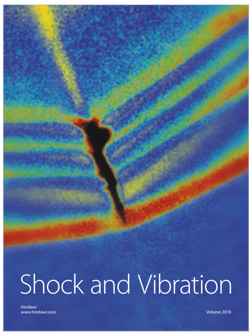

Modelling &
Simulation
in Engineering

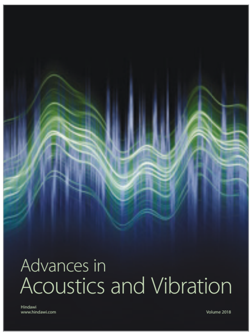
International Journal of
Aerospace
Engineering


International Journal of
Chemical Engineering


International Journal of
Antennas and
Propagation

Active and Passive
Electronic Components


Shock and Vibration


Advances in
Acoustics and Vibration