

Estimating computer network security scenarios with association rules

Anjani Kumar Singha [‡]

Department of Computer Applications

SRM Institute of Science and Technology (Deemed to be University)

Delhi-NCR Campus

Ghaziabad 201204

Uttar Pradesh

India

Harsh Pratap Singh [†]

Department of Computer Science and Engineering

Sri Vaishnav Vidhyapeeth Vishwavidyalaya

Indore 453331

Madhya Pradesh

India

Shakti Kundu [§]

School of Engineering and Technology

Computer Science Engineering

BML Munjal University

Gurugram 122413

Haryana

India

Pradeep Kumar Tiwari ^{*}

Department of Computer Science and Applications

Dr. Vishwanath Karad MIT World Peace University, Pune

Pune 411005

Maharashtra

India

Ajeet Singh Rajput [@]

Department of Computer Science & Engineering

Medicaps University

Indore

Madhya Pradesh

India

[‡] E-mail: anjani348@gmail.com

[†] E-mail: drharshprataps@gmail.com

[§] E-mail: shaktikundu@gmail.com

^{*} E-mail: Pradeeptiwari.mca@gmail.com (Corresponding Author)

[@] E-mail: ajeetsinghrajput@gmail.com

Abstract

Traditional NSSA (network security situational awareness) systems are problems familiarize to enormous for complicated network circumstances due to instrument limitations, inadequate and data fusion capabilities. This paper proposes the investigation of Association Rules (AR) based NSS (network security situation) prediction technology. Mining for regulations for associations to remedy this issue re- confidence support framework is enhanced by incorporating an interest. The assessment standard is revised, and the worth of AR is reassessed established on a discussion of associated standard ideas and algorithms AR mining. This research proposes an MFP-interest algorithm that combines an alert AR pattern with a degree of interest. The MFP-interest algorithm was evaluated. The researcher found that the MFP-interest algorithm is capable of accurately predicting NSS. The majority of time points have relative error ranges of less than 0.035.

Subject Classification: Primary 93A30, Secondary 49K15.

Keywords: Computer network security, Association rules, Data mining, Security scenarios, and Risk assessment.

1. Introduction

Network hazards, network infections, network infiltration, and network security issues are growing in severity as the use of all types of information networks continues to rise. Security flaws in networks, such as those caused by hackers, become increasingly apparent over time [1]. It's possible that relying solely on tried-and-true ways could backfire, wreaking havoc on users' privacy and financial security while undermining their faith in the network's reliability [2, 3]. Given the growing importance of network security, it's crucial to have a firm grasp of the state of the industry at any given time to protect against unforeseen dangers and maintain a safe connection environment. In light of these frequent network disasters, people have created sophisticated network protection measures such as intrusion detection technology and private network technology [4]. Each of these technologies has contributed to the evolution of network security, but they all have limitations and need to be revised [5]. Assessing and forecasting the NSS in real-time allows proactive network management before security problems occur, which helps prevent harm and maintains network resource security [6]. The rapid growth of Internet-related companies in China has outpaced the country's ability to implement adequate security measures. While the Internet simplifies data storage, transfer, and retrieval, there are still gaps in safety that prevent complete confidence in its use [7]. NSSA technology was developed out of this necessity [8]. The network data is prepared [9], abnormal data is searched

for, and the law of network anomalous data is summarized to evaluate if the data is possibly damaging to the network.

The main contributions of this research

- The application of association rule mining to intrusion detection, which has led to the development of new systems that can detect intrusions more accurately and efficiently than traditional methods.
- The development of anomaly detection systems that use association rule mining to identify anomalous behavior in network traffic.
- The development of attack prediction systems that use association rule mining to identify patterns in network traffic that are indicative of impending attacks.
- The development of risk assessment models that use association rule mining to estimate the risk of different security scenarios.

2. Related Work

Network security has been the subject of substantial study by academics and professionals in the United States and elsewhere, with findings categorized as either focusing on identifying misuse or anomaly detection [10]. Literature [11] shows that many researchers have committed time and energy to studying NSS, specifically two subtopics: NSS evaluation and NSS predictions. According to the results of the test, the hidden Markov model used to detect network service anomalies [12] in the literature has an excellent rate of recognition for network service anomalies. To combine many data sets into one, a Bayesian network-based scenario evaluation model was presented in the literature [13]. In the literature, a game-theory-based model for danger prediction and situational consciousness is proposed [14]. The software performs threat detection, then information fusion, and finally, a discrete hidden Markov game model predicts the intended assault intention. To forecast SVM (support vector machine), literature looks at the performance indices of different kernel functions [15]. Based on research into potential dangers, system administrators, and end-user actions, a hidden Markov gaming model is developed. Simulation results validate the model's ability to reflect security-related situations and inform administrators' choices constantly [16]. Based on the probable connection between the scenarios in the literature [17] constructs a Bayesian network evaluation architecture

and discusses the appropriate data broadcasting algorithm. At last, an illustration of the Bayesian network computation procedure is provided. The risks of the system are not restricted to these less damaging ways, and concealed compared to older tactics like viruses and website phishing. It complicates efforts to defend the network. Novel approaches and technologies for visualizing complicated network behavior at scale are described and studied in the literature [18], emphasizing the dynamics of networks and the unpredictability of network data. Data mining and multisensory data fusion are two important technologies in the literature [19]. To better assess the state of security, an organizational statistical assessment model is proposed in the literature [20]. The evaluation technique and accompanying calculation method in this statistical analysis are bottom-up, starting with the local level and working up to attack/vulnerability levels. To assess the susceptibility architecture, literature [21] established a LAN-based NSS evaluation system with two parts: network security danger level evaluation development trend prediction.

3. Research Method

3.1 *Establishment of NSSA System*

The NSSA System is Established (3.1). To identify potential threats to a network's security, NSSA requires an in-depth understanding of the external environment's dynamics [22]. Using big data as its foundational technological advances, NSSA is an efficient means of detecting, investigating, and counteracting network security risks items mining is the technique of mechanically identifying and removing possibly dangerous items from a vast volume of data stored on a network of computers. Possibly harmful network information can be identified through data preparation, network searches, and law of aberrant networking data summarization. The state of network safety now defies a simple qualitative evaluation. In order minimize the number of network threats, uncover hidden networking dangers immediately, and offer a solid platform for anticipating networking security trends, an acceptable, effective, and reliable qualitative evaluation of the present-day network security situation in real time is required. Create a method to resolve problems that will help you evaluate the security situation quantitatively. In order to categories, define, calculate, and analyses the safety metrics, the National Social Security Administration system generates the NSSA model. The security situational awareness model is adaptive and can learn

from past events to improve in the future. A dispersed multilayer system structure, consisting primarily of an information gathering terminals and an analysis center, is recommended in light of the macro network's massive size and data volume. Figure 1 displays the structure of the system's hierarchical diagram.

3.2 Computer NSS Prediction

The goal of association rule mining is to discover hidden connections in data. Elements selected from a vast data set. People are becoming increasingly interested in data mining as the volume of database information grows.

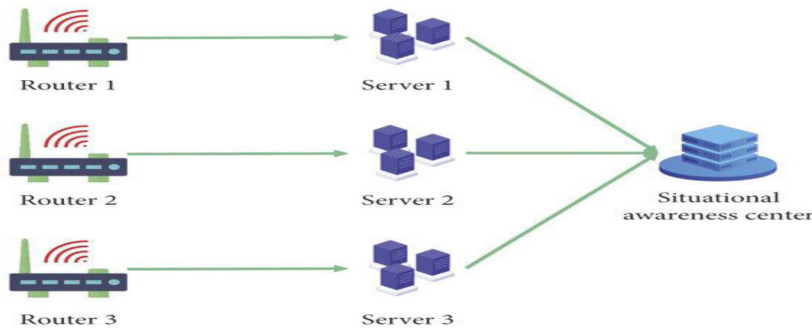


Figure 1

Diagram of the system structure hierarchy

Mining augmented reality consists of two main stages:

Sets of items that occur frequently (1). Utilize the user-supplied minimum support value min_supp to uncover all frequently occurring item sets Create AR

(2). There is at least AR min_conf in each maximum frequent item set, where min_conf is the minimal confidence level the user specified.

The exact terminology for this case is

$$\text{confidence}(C \Rightarrow D) = P(D \cup C), \text{support}(C \Rightarrow D) = P(C \cup D), \quad (1)$$

In the first of these two phases, we solve the basic problem of AR mining and benchmark the performance of AR mining algorithms by discovering all frequent item sets as rapidly and efficiently as possible. The answer to the second stage is simple. Most existing AR mining techniques are proposed for the first step's sub problem. For the purposes of numerical thinking and estimation, getting knowledge of what is

happening in the field of security in networks compared to large-scale sources of information [23]. The nature and precision of quantitative perceptions are directly impacted by the elements and outcomes that are gathered. There is a dearth of relevant academic literature because this area of study is still in its infancy [24]. The framework for further study in this area was provided by earlier work on feature identification, classification analysis, and clusters. In, Figure 2 depicts the three stages in the process flow of the AR mining-based computer network security situation prediction technique developed in this paper. Virus data gathered from end users is gathered using the fractional interval spectrum feature extraction method. To get the qualitative recursion entropy ratio $S_n(r, i)$ of the underlying NSS, we first calculate the cluster d_{ij} between its features points for a calculated $M(i)$, single-component NSS time sampling series Y_i .

$$S_n(r, i) = \frac{M(i)}{M - (n-1)r} \quad (2)$$

The size of the NSS sequence's vector field encapsulation is calculated. The NSS channel equalization model during conditions of feature disturbance is as follows:

$$NSS_AV = \lim_{N \rightarrow \infty} [MN(r) - MNn + 1(r)]. \quad (3)$$

When an alarm occurs, it is recorded in the alert logs of various network safety components. Naturally, alarm logging is a multimodal record, and the AR of the alert log must be assigned to my multimodal AR in the careful log databases. In this chapter, we expand the method known as FP-growth as part of the MFP-growth (multifaceted FP-growth) method for mine multidimensional AR [24], and we further refine the AR template along with interest degree to apply constraints to multidimensional AR mining using the MFP-interest technique.

For AR $X \Rightarrow Y$, support is defined as

$$\text{support}(A \Rightarrow B) = P(A \cup B). \quad (4)$$

Confidence is defined as

$$\text{confidence}(A \Rightarrow B) = \frac{P(A \cup B)}{P(A)} \quad (5)$$

Although employing the minimal support criterion and the minimum trust level can assist in eliminating or decreasing mining useless rules, the AR that meets both thresholds becomes a powerful AR.

The meaning of interest is

$$\text{interest } X \Rightarrow Y = \frac{P(A \cup B)}{P(A) \cdot P(B)} \quad (6)$$

The AR degree that is relevant evaluates the functionality, novelty, simplicity, and assurance of rules. As a result, the required minimum support establishment must be as low as possible during the AR mining process, which places greater demands on the efficiency when the minimum encouragement establishing is common, the usability of AR mining is significantly hampered by the increase of a large number of useless rules; an AR template can restrict the types of AR generated, produce only those regulations that users are interested in, and significantly decrease. The MFP-algorithm proposed in this paper operates according to the abovementioned conditions. The MFP-interest algorithm implements the divide-and-conquer strategy and utilizes the FP tree, a compacted data storage structure that is a component of the MFP-growth algorithm: Condense the alarm log library containing frequently used item sets into an FP-tree representing frequent patterns, preserving the item sets' count of frequency and dimensional attributes. Subsequently, the compacted database is partitioned into a collection of conditional databases. An AR template is employed to regulate the expansion of regular prints; pattern growth is restricted to those patterns that meet the requirements of the template, and the support of the template is computed. Data mining establishes the NSSA threshold, and only data that shares network security situation characteristics are chosen for mining potentially hazardous data in the network to the fullest extent feasible to prevent the erroneous perception that can arise from similarity[9]. The AR mining-based computer network security situation prediction technology has been fully developed thus far.

4. Results Analysis and Discussion

It constitutes a collaborative and public operating system on the web. The network state enables the transmission and sharing of information and data, satisfying the need for knowledge resources among individuals and decreasing the expense associated with their procurement. Although the computer network facilitates work and daily life for individuals, it also presents criminals with possibilities. Regardless of the type of malicious attack, cybersecurity will always jeopardize the computer network. Consequently, hackers have emerged as the predominant threat to

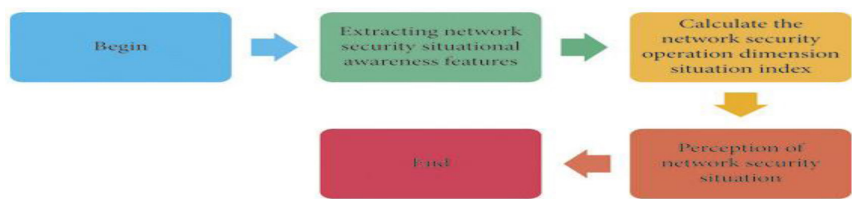


Figure 2

Flowchart of an AR mining-based computer NSS prediction system

computer network security in the big data era. In the era of big data, which is characterized by fraudulent data theft, computer users must increase their network safety awareness to prevent financial losses. To prevent unauthorized access to personal information facilitated by malicious actors employing virus programs, customers ought to exercise caution when connecting their machines to networks, refrain from randomly clicking on dangerous web links, and refrain from uploading sensitive data without verifying its security.

Additionally, to safeguard users’ devices against malicious intrusion by criminals, users should configure access permissions for the home network. It is strongly advised that users seek the aid of qualified network upkeep personnel without delay to rectify any issues with the computer system, such as password leakage or abnormal operation. In this investigation, the running times of frequently occurring item sets mined for alarm log datasets of varying sizes using MFP-interest algorithm, MFP-growth algorithm, and the MApriori algorithm are compared. The comparison of the efficiency of smaller-scale alarm log mining algorithms is illustrated in Figure 3. Figure 3 demonstrates that the MFP-growth and

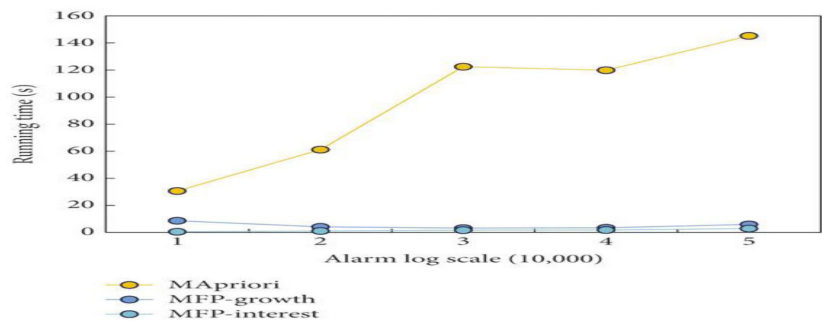


Figure 3

A comparison of the performance of algorithms for small-scale alerting log mining

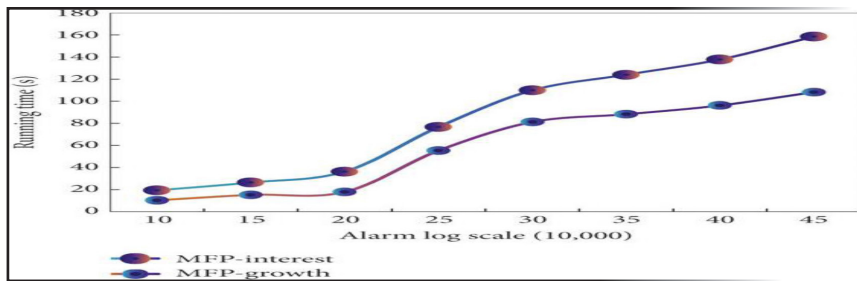


Figure 4

An analysis of the mining time of enormous alarm records

MFP-interest algorithms exhibit significantly higher efficiency than the MAPriori algorithm. As a result, the MAPriori algorithm could be more suitable for the mining of immense alarm records. Through manual authorization, the network's infrastructure details are incorporated into the resource's catalog and subsequently merged with the data gathered by the person who collected the data about the network, which are subsequently loaded into the network security fundamental database.

The record of network security events is constructed through the process of correlation evaluation. The outcomes of the analysis constitute a fundamental fuzzy matrix. The network security situation is assessed by integrating the fuzzy matrix and the analytical structure procedure. The evaluation's results and entire specifics are presented after the assessment. Various collectors of data process and store the original data in the database. As part of the NSS evaluation process, by analyzing the initial data, a complete assessment of the entire NSS is achieved, and security incidents are addressed. Meanwhile, by selecting a node, one can perceive the state of network security and compute the present rating of the neighboring network. Frequently, alarm logs are vast, and mining algorithms are confronted with greater performance demands when faced with enormous alarm logs. When the threshold for minimum support is set to a low value, the MAPriori algorithm cannot complete the mining task within the allotted time due to the large-scale alarm log.

An analysis of the execution times of the MFP-growth and MFP-interest algorithms for mining sets of commonly used items is presented in Figure 4. Consider Figure 5. As the size of alarm records expands, there is a corresponding rise in the quantity. The amount of item routine sets produced by the MFP-growth algorithm, in the absence of relationship template constraints, is significantly greater than that of the MFP-interest algorithm. After being processed, a substantial quantity of alarm

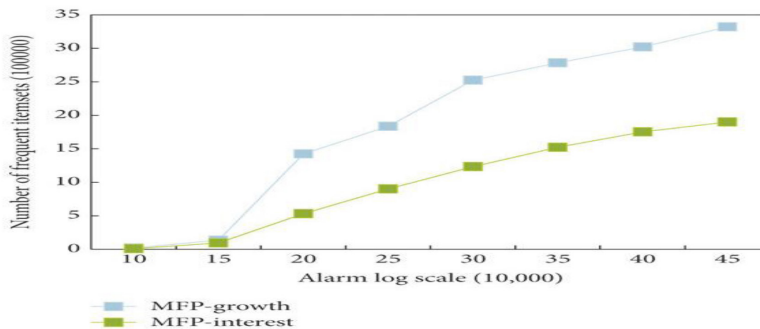


Figure 5

Typical Item Set Mining: A Comparison

intelligence gathered by data gatherers is entered into the fundamental database. The security event information must be analyzed to eliminate redundant alarm information.

The data presented in the alarm is intricately linked to the potential security breaches occurring within the network. The network management evaluation entails the examination of the security of the network alarm, as well as the classification, delouse, and analysis of the alarm data. Assembly analysis requires the following steps:

- Perform association evaluation on network security alarm data.
- Invoke the similarity function.
- Process secure event data.
- Could you keep it in a safety database?
- Generate statistically significant outcomes.

In comparison to the other two models for prediction (ref. [17] and ref. [19]), the accuracy of forecasts can be enhanced by this approach. The precise data are illustrated in Figure 6. As illustrated in Figure 6, the approach to predicting the fact that is the most precise is the one that is proposed in this paper; all other methods for predicting exhibit flaws. The data obtained during the primary data collection phase of the NSS evaluation module are subjected to a statistical analysis utilizing the statistically analyzed information and the key analysis secured event a database that is generated by the statistical analysis process. NSS evaluation is then executed through the calculation module. Upon the detection of an

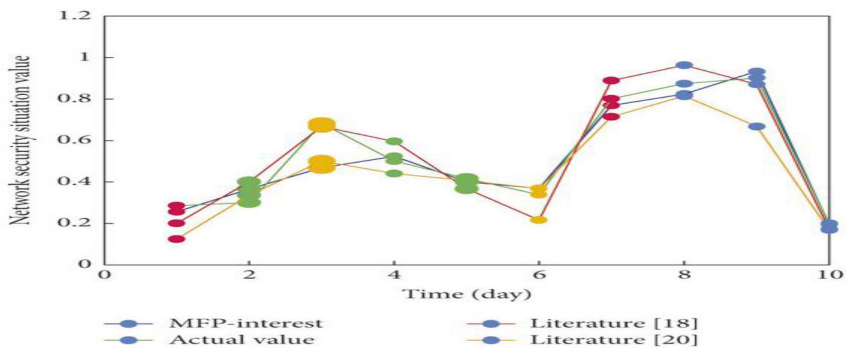


Figure 6

Evaluation of several approaches of forecasting in terms of NSS values

imminent danger within the system, an established reaction strategy is implemented.

Following the implementation of the internal response measures derived from the results of the NSS assessments, the system notifies the network's protection administrator of any system threats. For the plug-in to invoke the linkage response, it is necessary to analyze and compare response times, implement suitable measures, and transmit an email.

5. Conclusion

In summary, with the official advent of the era of big data, there is a growing concern among individuals regarding the security of computer networks. Users will endure irreparable harm if a substantial quantity of critical data is misplaced or stolen from their machine's network. Using the study of computer network security situation prediction technological advances based on AR extraction of minerals, this paper demonstrates that the network security situation precision during design technology is substantially greater than that of conventional technology, thereby providing a framework to promote network security situation data integration and knowledgeable evaluation and providing secure network managers with a real-time and dependable management choice basis. Compared to conventional prediction methods, the MFP-growth algorithm exhibits superior nonlinear operation and mapping capabilities, approximates functional interactions with arbitrary precision, demonstrates robust adaptability, and achieves higher accuracy in network security situation prediction. The findings of a situation assessment influence NSS forecasts. If the situation assessment yields substantial

deviations, those results will undeniably affect the network security situation prediction outcomes. Predicting the index data is the initial step; utilizing predicted index data to assess the anticipated situation value is the subsequent step.

References

- [1] H. Zhang, C. Kang, and Y. Xiao, "Research on network security situation awareness based on the LSTM-DT model," *Sensors*, vol. 21, no. 14, p. 4788 (2021).
- [2] A. Dureja, and P. Pahwa, "A comparative approach for classifying retinal OCT images based on deep learning framework," *Journal of Discrete Mathematical Sciences and Cryptography*, 25(4), 859-870 (2022).
- [3] J. Sharma, Sandeep Kumar, Kumar, Anil, Ashtagi, Rashmi & Jain, Rekha. OCA: An intelligent model for improving security breach of biometric based authentication systems, *Journal of Discrete Mathematical Sciences and Cryptography*, 26:5, 1415–1425 (2023), DOI: 10.47974/JDMSC-1765.
- [4] A. K. Singha, S. Zubair, A. Malibari, N. Pathak, S. Urooj, and N. Sharma, "Design of ANN Based Non-Linear Network Using Interconnection of Parallel Processor," *Computer Systems Science & Engineering*, 46(3) (2023).
- [5] S. Zubair, A. K. Singha, N. Pathak, N. Sharma, S. Urooj, and S. R. Larguech, "Performance Enhancement of Adaptive Neural Networks Based on Learning Rate," *Computers, Materials & Continua*, 74(1) (2023).
- [6] N. Pathak, S. T. Siddiqui, A. K. Singha, H. G. Mohamed, S. Urooj, and A. R. Patil, "Smart quarantine environment privacy through IoT gadgets using blockchain," *Intell Autom Soft Comput* (Accepted) (2023).
- [7] G. Ramesh, A. Sharma, D. V. Lalitha Parameswari, C. Mallikarjuna Rao, and J. Somasekar, "Blockchain in healthcare: Moving towards a methodological framework for protecting Biomedical Databases," *Journal of Discrete Mathematical Sciences and Cryptography*, 25(4), 891-901 (2022).
- [8] J. Kong, C. Yang, J. Wang et al., "Deep-stacking network approach by multisource data mining for hazardous risk identification in IoT-based intelligent food management systems," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1194565, 16 pages (2021).

- [9] X.-B. Jin, W.-Z. Zheng, J.-L. Kong et al., "Deep-learning temporal predictor via bidirectional self-attentive encoder- decoder framework for IOT-based environmental sensing in intelligent greenhouse," *Agriculture*, vol. 11, no. 8, p. 802 (2021).
- [10] A. K. Singha, and S. Zubair , " Combination of Optimization Methods in a Multistage Approach for a Deep Neural Network Model," *International Journal of Information Technology*, 1-7 (2023).
- [11] Z. Ran, L. Min, Y. Yifeng, Z. Qikun, and C. Zengyu, "Pre- diction algorithm for network security situation based on BP neural network optimized by SA-SOA," *International Journal of Performability Engineering*, vol. 16, no. 8, p. 1171 (2020).
- [12] Q. Liu and M. Zeng, "Network security situation detection of internet of things for smart city based on fuzzy neural net- work," *International Journal of Reasoning-Based Intelligent Systems*, vol. 12, no. 3, p. 222 (2020).
- [13] P. Lin and Y. Chen, "Network security situation assessment based on text SimHash in big data environment," *International Journal on Network Security*, vol. 21, no. 4, pp. 699–708 (2019).
- [14] Y. Zhu and Z. Du, "Research on the key technologies of network security-oriented situation prediction," *Scientific Programming*, vol. 2021, Article ID 5527746, 10 pages (2021).
- [15] A. K. Singha, N. Pathak, N. Sharma, P. K. Tiwari, and J. P. C. Joel , " COVID-19 Disease Classification Model Using Deep Dense Convolutional Neural Networks," In *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2022, Volume 2* (pp. 671-682) (2022).
- [16] N. Ta, H. Li, S. Liu, and Y. Zuo, "Mining key regulators of cell re- programming and prediction research based on deep learning neural networks," *IEEE Access*, vol. 8 (2020).
- [17] X. Qiao, "Study on the computer network security evaluation based on GABP algorithm," *Boletin Tecnico/Technical Bulletin*, vol. 55, no. 18, pp. 497–504 (2017).
- [18] F. Catal, N. Tcholtchev, E. Hořfig, and A. Hoffmann, "Visu- alization of traffic flows in a simulated network environment to investigate abnormal network behavior in complex network infrastructures," *Procedia Computer Science*, vol. 151, pp. 279–286 (2019).

- [19] C. Yang, A. Mohammadi, and Q.-W. Chen, "Multi-sensor fusion with interaction multiple model and chi-square test tolerant filter," *Sensors*, vol. 16, no. 11, p. 1835 (2016).
- [20] A. K. Singha, N. Pathak, N. Sharma, P. K. Tiwari, and J. P. C. Joel, "Forecasting COVID-19 Confirmed Cases in China Using an Optimization Method," In *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2022, Volume 2* (pp. 683-695) (2022).
- [21] R. Hou, G. Ren, C. Zhou, H. Yue, H. Liu, and J. Liu, "Analysis and research on network security and privacy security in ubiquitous electricity Internet of Things," *Computer Communications*, vol. 158, pp. 64–72 (2020).
- [22] S. Yang, D. Yin, X. Song et al., "Security situation assessment for massive MIMO systems for 5G communications," *Future Generation Computer Systems*, vol. 98, pp. 25–34 (2019).
- [23] S. T. Siddiqui, A. K. Singha, M. O. Ahmad, M. Khamruddin, and R. Ahmad, "IoT devices for detecting and machine learning for predicting COVID-19 outbreak," In *Recent Trends in Communication and Intelligent Systems: Proceedings of ICRTCIS 2021* (pp. 107-114) (2022).
- [24] J.P. Wei and Z. Zhou, "Research on security of information sharing in Internet of Things based on key algorithm," *Future Generation Computer Systems*, vol. 88, no. NOV, pp. 599–605 (2018).