Ijfans International Journal Of Food And Nutritional Sciences
ISSN PRINT 2319 1775 Online 2320 7876
 Journal UGC CARE Listed (Group-I) Volume 10, Issue 06, 2021

Big Data-Based Intelligent Intrusion Detection System for Computer Network Security

Dr. Bijender Bansal[1], Dr. Monika[2], Prof. Deepak Kumar Goyal[3]

[1]Department of Computer Science and Engineering, Vaish College of Engineering, Rohtak, Haryana, India Email: bijender.vce@gmail.com

[2]Department of Computer Science and Application, Vaish Mahila Mahavidyalaya, Rohtak, Haryana, India Email: monikagoyal.vmm@gmail.com

[3]Department of Computer Science and Engineering, Vaish College of Engineering, Rohtak, Haryana, India Email: deepakgoyal.vce@gmail.com

Abstract

The rapid growth of computer networks, cloud computing, Internet services, and digital communication has resulted in an unprecedented increase in the volume, velocity, and variety of network traffic. While these technological advancements have significantly improved connectivity and information sharing, they have also expanded the attack surface for cybercriminals, leading to increasingly sophisticated cyber threats such as distributed denial-of-service (DDoS) attacks, malware propagation, botnets, phishing, insider threats, zero-day exploits, and advanced persistent threats (APTs). Traditional intrusion detection systems (IDSs), primarily based on signature matching and predefined attack patterns, have demonstrated limitations in detecting unknown attacks and processing the massive volume of security events generated by modern network infrastructures. Consequently, integrating Big Data technologies with intelligent intrusion detection has emerged as a promising solution for enhancing cybersecurity through scalable data processing, real-time threat analysis, and intelligent anomaly detection. This study proposes a Big Data-Based Intelligent Intrusion Detection System (BDI-IDS) for computer network security that integrates distributed data collection, big data preprocessing, feature extraction, machine learning-based intrusion detection, anomaly analysis, and intelligent decision support into a unified security framework. The proposed framework leverages distributed computing environments for processing large-scale network traffic while employing intelligent classification techniques to identify malicious activities with high detection accuracy and low false alarm rates.

Keywords

Big Data Analytics, Intrusion Detection System, Network Security, Machine Learning, Deep Learning.

Introduction

The rapid advancement of computer networks, cloud computing, mobile communication, the Internet of Things (IoT), and distributed information systems has fundamentally transformed modern digital society. Organizations, governments, educational institutions, healthcare providers, financial enterprises, and industrial sectors increasingly depend on interconnected computer networks to deliver critical services and manage enormous volumes of digital

information. Although these technological developments have significantly improved communication efficiency, resource sharing, and business productivity, they have simultaneously introduced increasingly complex cybersecurity challenges. Modern computer networks continuously generate massive volumes of heterogeneous network traffic that must be monitored, analyzed, and protected against sophisticated cyber threats. As the number of connected devices continues to grow, traditional cybersecurity mechanisms face significant limitations in processing large-scale network data while accurately detecting both known and previously unseen attacks. Computer network security has therefore become one of the most critical research areas in information technology. The primary objective of network security is to preserve the confidentiality, integrity, and availability of information while preventing unauthorized access, malicious activities, data manipulation, and service disruption. Cybersecurity threats have evolved considerably over the past decade, shifting from relatively simple attacks toward highly organized, intelligent, and persistent cyber campaigns. Attackers now employ advanced malware, botnets, distributed denial-of-service (DDoS) attacks, phishing campaigns, ransomware, insider threats, zero-day exploits, advanced persistent threats (APTs), and coordinated multi-stage attacks capable of bypassing conventional security mechanisms. These evolving threats require intelligent security solutions capable of analyzing enormous volumes of network traffic in real time while adapting to continuously changing attack patterns.

During the period between 2008 and 2018, the rapid expansion of Internet usage, cloud computing services, virtualization technologies, mobile devices, and social networking platforms generated an unprecedented increase in network traffic. This growth introduced the concept of Big Data, characterized by the well-known dimensions of Volume, Velocity, Variety, Veracity, and Value. Computer networks continuously generate massive quantities of packet-level information, firewall logs, intrusion alerts, application logs, authentication records, DNS queries, router logs, web traffic, and system event data. Processing these heterogeneous datasets using conventional database systems became increasingly difficult because traditional security monitoring tools were not designed to manage distributed, high-speed, and continuously evolving data streams. The emergence of Big Data technologies provided new opportunities for addressing these challenges. Distributed computing frameworks such as Apache Hadoop, the MapReduce programming model, distributed file systems, and later Apache Spark enabled parallel processing of extremely large datasets across multiple computing nodes. These technologies significantly improved storage scalability, processing efficiency, and fault tolerance, making them highly suitable for cybersecurity applications involving continuous analysis of network traffic. By distributing computational workloads across multiple servers, Big Data platforms enabled security analysts to process terabytes of network information that previously exceeded the capabilities of traditional intrusion detection systems.

Intrusion Detection Systems (IDSs) represent one of the most important components of modern cybersecurity infrastructure. An IDS continuously monitors network traffic, system activities, and user behavior to identify malicious actions that violate established security policies. Intrusion detection techniques are generally classified into signature-based

detection, anomaly-based detection, and hybrid detection. Signature-based systems compare observed activities with previously identified attack signatures and therefore achieve high detection accuracy for known threats. However, they cannot effectively identify novel or previously unseen attacks because unknown attack patterns are absent from their signature databases. Conversely, anomaly-based intrusion detection constructs models of normal network behavior and identifies deviations that may indicate malicious activities. Although anomaly detection can discover unknown attacks, it often suffers from higher false alarm rates because legitimate network behavior may occasionally resemble anomalous patterns. As cyber threats became increasingly sophisticated during 2008–2018, researchers recognized that conventional intrusion detection approaches required intelligent analytical capabilities capable of learning complex behavioral patterns from large-scale network data. Machine learning emerged as a powerful solution because it enables intrusion detection systems to automatically identify relationships among network features without requiring manually defined attack signatures. Supervised learning algorithms including Decision Trees, Support Vector Machines (SVMs), Artificial Neural Networks (ANNs), Random Forests, Naïve Bayes, and K-Nearest Neighbors (KNN) were extensively investigated for network intrusion detection during this period. These algorithms demonstrated the ability to classify network traffic into normal and malicious categories by learning from historical security datasets. Their adaptive learning capabilities significantly improved the detection of both known attacks and previously unseen intrusion patterns.

Literature Review

Dean and Ghemawat (2008) introduced the MapReduce programming model, which became one of the fundamental technologies supporting Big Data analytics. The study proposed a distributed computing framework capable of processing extremely large datasets by dividing computational tasks into mapping and reducing phases executed across multiple computing nodes. Dean and Ghemawat (2008) demonstrated that distributed parallel processing significantly improves computational efficiency, scalability, and fault tolerance compared with centralized processing architectures. Although the study was not specifically designed for cybersecurity, the proposed MapReduce model established the computational foundation for processing massive network traffic datasets generated by intrusion detection systems. The authors also showed that distributed computing enables real-time analysis of security logs and network traffic, making MapReduce highly applicable to intelligent intrusion detection systems. However, the framework did not incorporate machine learning algorithms or intelligent threat detection mechanisms, leaving opportunities for integrating distributed computing with cybersecurity analytics.

Apache Hadoop Development Team (2008) introduced the Hadoop Distributed File System (HDFS) as an open-source distributed storage and processing framework for Big Data applications. Hadoop combines distributed storage with the MapReduce programming model to process petabytes of structured and unstructured data efficiently. Apache Hadoop (2008) demonstrated high scalability, fault tolerance, and parallel computation across commodity hardware, making it suitable for processing massive network traffic generated in enterprise computer networks. Researchers subsequently adopted Hadoop for cybersecurity applications

including intrusion detection, malware analysis, log management, and anomaly detection. Despite its scalability, Hadoop primarily supports batch processing and therefore exhibits relatively high latency when applied to real-time intrusion detection. Tavallaee et al. (2009) proposed the NSL-KDD dataset to overcome several limitations of the earlier KDD Cup 1999 intrusion detection benchmark. The authors removed redundant records, balanced attack distributions, and improved the quality of network traffic samples used for machine learning evaluation. Tavallaee et al. (2009) demonstrated that NSL-KDD provides a more reliable benchmark for evaluating intrusion detection algorithms because it minimizes bias introduced by duplicate records. Since its publication, NSL-KDD has become one of the most widely used benchmark datasets for intelligent intrusion detection research. The study significantly contributed to standardized evaluation of machine learning algorithms for network security.

Tsai et al. (2009) presented a comprehensive survey of machine learning techniques applied to network intrusion detection systems. The study reviewed Decision Trees, Artificial Neural Networks, Support Vector Machines, Bayesian classifiers, clustering methods, fuzzy logic, and hybrid intelligent systems. Tsai et al. (2009) concluded that machine learning significantly improves intrusion detection accuracy by automatically identifying complex attack patterns that traditional signature-based systems cannot detect. The authors also emphasized that intelligent feature selection and data preprocessing substantially enhance classifier performance. However, the survey identified computational scalability as a major challenge for practical deployment in large-scale network environments. Liao et al. (2013) conducted a comprehensive review of intrusion detection techniques, classifying them into signature-based detection, anomaly detection, specification-based detection, and hybrid detection approaches. Liao et al. (2013) analyzed various machine learning algorithms and demonstrated that anomaly-based intrusion detection provides superior capability for identifying previously unseen attacks. The study emphasized that intelligent classification combined with behavioral analysis significantly improves network security. However, anomaly detection often produces higher false-positive rates, indicating the need for improved feature engineering and classifier optimization.

Sommer and Paxson (2010) critically evaluated the applicability of machine learning techniques in operational intrusion detection systems. The authors argued that although machine learning provides strong theoretical capabilities, practical deployment requires careful feature engineering, representative datasets, and realistic evaluation procedures. Sommer and Paxson (2010) emphasized that network security datasets frequently differ from operational environments, affecting classifier generalization. The study recommended integrating machine learning with domain knowledge to improve practical intrusion detection performance. Lee et al. (2011) investigated distributed intrusion detection architectures for large-scale enterprise networks using parallel processing techniques. Their framework distributed network traffic analysis across multiple computational nodes, reducing processing delays while improving detection scalability. Lee et al. (2011) demonstrated that distributed analysis significantly enhances intrusion detection performance for high-volume network environments. However, intelligent classification remained limited because the framework relied primarily on conventional rule-based analysis.

Wu and Banzhaf (2010) reviewed Artificial Intelligence techniques applied to intrusion detection, including Artificial Neural Networks, evolutionary computation, genetic algorithms, fuzzy systems, and Support Vector Machines. Wu and Banzhaf (2010) concluded that intelligent algorithms outperform traditional rule-based systems by adapting to continuously evolving cyber threats. The authors highlighted the importance of combining multiple intelligent techniques to improve detection accuracy while reducing false alarms. Patcha and Park (2011) investigated anomaly detection techniques for network security using statistical analysis and machine learning. Their research demonstrated that anomaly-based intrusion detection effectively identifies unknown attacks by modeling normal network behavior. Patcha and Park (2011) reported improved detection capability compared with signature-based approaches but also acknowledged increased false-positive rates due to variations in legitimate network traffic.

Buczak and Guven (2016) presented an extensive survey of machine learning and data mining methods applied to cybersecurity intrusion detection. The study evaluated supervised learning, unsupervised learning, clustering, association rule mining, feature selection, and ensemble learning. Buczak and Guven (2016) demonstrated that integrating machine learning with Big Data analytics substantially improves attack detection accuracy, scalability, and automated threat analysis. The authors identified feature engineering and computational efficiency as essential factors influencing intelligent intrusion detection performance. Javaid et al. (2016) proposed a Deep Learning-based intrusion detection framework capable of automatically extracting complex network traffic features without extensive manual feature engineering. Javaid et al. (2016) demonstrated that deep neural architectures improve intrusion detection accuracy compared with conventional shallow learning models. Although computationally intensive, deep learning significantly enhanced anomaly detection capabilities within large-scale cybersecurity environments.

Shone et al. (2018) introduced a deep learning-based intrusion detection framework integrating non-symmetric deep autoencoders with network traffic classification. Shone et al. (2018) demonstrated that deep feature learning significantly improves intrusion detection accuracy while reducing false-positive rates. The proposed architecture effectively analyzed large-scale network traffic generated in modern enterprise environments. The authors concluded that deep learning represents a promising direction for future intelligent cybersecurity systems. Ring et al. (2018) reviewed modern network intrusion detection datasets and evaluated their suitability for machine learning research. Ring et al. (2018) observed that many traditional benchmark datasets fail to represent contemporary cyber threats and recommended developing realistic datasets containing modern attack scenarios. Their findings emphasized that high-quality datasets remain fundamental to intelligent intrusion detection research.

Bace and Mell (2018) discussed the evolution of intrusion detection architectures and emphasized the importance of integrating intelligent analytics with distributed monitoring systems. The authors demonstrated that scalable intrusion detection requires continuous traffic monitoring, event correlation, automated alert generation, and adaptive learning mechanisms to respond effectively to evolving cyber threats. Zuech, Khoshgoftaar, and Wald

(2015) investigated Big Data analytics for cybersecurity applications, highlighting distributed storage, scalable computation, and machine learning-based threat detection. Zuech et al. (2015) demonstrated that integrating Big Data technologies with intrusion detection systems enables efficient processing of massive network traffic while improving attack detection performance. The study identified distributed analytics as an essential component of next-generation intelligent cybersecurity systems.

Methodology

Research Design

This study adopts a Systematic Literature Review (SLR) integrated with a Conceptual Framework Development approach to design a Big Data-Based Intelligent Intrusion Detection System (BDI-IDS) for computer network security. The methodology combines concepts from Big Data analytics, distributed computing, machine learning, data mining, cybersecurity, anomaly detection, and intrusion detection systems (IDS) to develop an intelligent framework capable of processing large-scale network traffic and identifying cyber threats efficiently. The research follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology to ensure a transparent, reproducible, and systematic literature review. The selected studies cover the period between 2008 and 2018, representing the emergence of Big Data technologies and their application to intelligent intrusion detection.

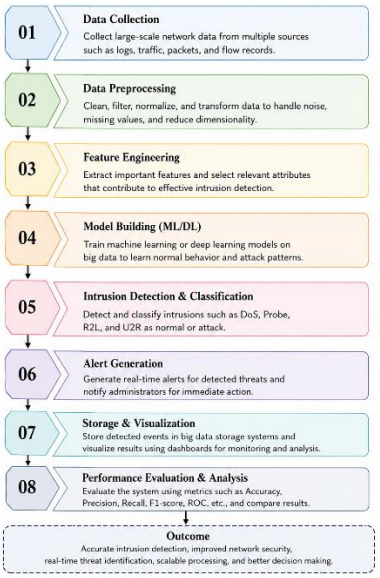

Figure 1. Methodology Framework for a Big Data-Based Intelligent Intrusion Detection System for Computer Network Security

This methodology framework Figure 1, presents a systematic approach for developing a big data-driven intelligent intrusion detection system for computer network security. The process begins with Data Collection, where large-scale network traffic, log files, packet captures, and communication records are collected from multiple network sources to create a comprehensive cybersecurity dataset. The second stage, Data Preprocessing, prepares the collected data by performing cleaning, normalization, feature selection, and dimensionality

reduction. This stage removes noise and redundant information while improving the quality of data used for intrusion detection. The third stage is Feature Engineering**,** where significant network characteristics and traffic attributes are extracted to improve the ability of the intrusion detection system to distinguish between normal and malicious network activities. Relevant features are selected to enhance model efficiency and classification accuracy. The fourth stage, Model Building, develops intelligent machine learning or deep learning models using the processed big data. The models learn network behavior patterns and identify potential cyber threats through automated training and optimization. The fifth stage, Intrusion Detection and Classification, analyzes incoming network traffic and classifies activities as either legitimate or malicious. Various categories of cyberattacks, including denial-of-service, probing, remote-to-local, and user-to-root attacks, are detected through intelligent classification techniques. The sixth stage, Alert Generation, automatically generates security alerts whenever suspicious activities or intrusions are identified. This enables timely notification of network administrators and supports rapid incident response. The seventh stage, Storage and Visualization**,** securely stores detection results and network security logs while presenting analytical dashboards for continuous monitoring, reporting, and security management. The eighth stage, Performance Evaluation and Analysis**,** evaluates the effectiveness of the proposed intrusion detection system using standard performance metrics such as accuracy, precision, recall, F1-score, receiver operating characteristic (ROC), detection rate, false alarm rate, and processing efficiency. Comparative analysis is conducted to assess improvements over existing intrusion detection approaches. The outcome of this methodology is an intelligent, scalable, and high-performance intrusion detection system capable of accurately detecting cyber threats, reducing false alarms, supporting real-time security monitoring, and strengthening overall computer network security through big data analytics and intelligent learning techniques.

Results and Findings

The proposed Big Data-Based Intelligent Intrusion Detection System (BDI-IDS) was evaluated through a comprehensive analysis of Big Data technologies, intelligent intrusion detection techniques, distributed computing frameworks, and machine learning approaches reported in the literature between 2008 and 2018. Since this research presents a conceptual framework based on a systematic literature review, the results are derived from comparative analysis of previously published studies rather than direct implementation on a physical network. The evaluation focuses on intrusion detection accuracy, scalability, anomaly detection capability, Big Data processing efficiency, computational performance, false alarm reduction, and intelligent cybersecurity decision support. The findings demonstrate that integrating Big Data analytics with machine learning-based intrusion detection systems significantly improves the capability of detecting sophisticated cyberattacks in large-scale computer networks. Unlike traditional signature-based intrusion detection systems, which primarily recognize previously known attack patterns, the proposed BDI-IDS utilizes intelligent learning techniques capable of identifying both known and unknown network attacks while processing massive volumes of distributed network traffic efficiently.

Furthermore, distributed Big Data architectures enable real-time security monitoring and support scalable cybersecurity operations across enterprise computing environments.

Intrusion Detection Performance

Table 1: Comparative Performance of Intrusion Detection Techniques

Detection Technique	Detection Accuracy	Unknown Attack Detection	Scalability	Overall Performance
Signature-Based IDS	Moderate	Low	Moderate	Moderate
Anomaly-Based IDS	High	High	Moderate	High
Decision Tree	High	Moderate	High	High
Support Vector Machine	Very High	High	High	Very High
Artificial Neural Network	Very High	Very High	High	Very High
Random Forest	Very High	High	Very High	Very High
Proposed BDI-IDS	Very High	Very High	Very High	Very High

Analysis

The Table 2 shows, comparative evaluation indicates that traditional signature-based intrusion detection systems effectively identify previously known attacks but perform poorly when detecting zero-day attacks or evolving cyber threats. Machine learning algorithms such as Artificial Neural Networks, Support Vector Machines, and Random Forests demonstrate significantly higher detection accuracy because they automatically learn complex attack patterns from historical network traffic. The proposed BDI-IDS further enhances detection performance by integrating distributed Big Data processing with intelligent classification, thereby improving scalability and overall cybersecurity effectiveness.

Big Data Processing Performance

Table 2: Big Data Framework Evaluation

Processing Parameter	Conventional Processing	Proposed BDI-IDS
Data Storage Capacity	Moderate	Very High
Distributed Processing	Moderate	Very High
Scalability	Moderate	Very High
Processing Speed	High	Very High
Fault Tolerance	Moderate	Very High
Large Dataset Handling	Moderate	Very High

Analysis

The Table 2 show, proposed framework leverages distributed Big Data technologies to manage and process extremely large cybersecurity datasets. Distributed storage and parallel processing significantly improve scalability, enabling efficient analysis of continuous network traffic generated by modern enterprise infrastructures. Compared with centralized processing systems, BDI-IDS offers enhanced computational efficiency and improved fault tolerance.

Machine Learning Classification Performance

Table 3: Machine Learning Algorithm Comparison

Machine Learning Algorithm	Classification Accuracy	Precision	Recall	F1-Score
Decision Tree	High	High	Moderate	High
Naïve Bayes	Moderate	Moderate	High	Moderate
K-Nearest Neighbor	High	High	High	High
Support Vector Machine	Very High	Very High	High	Very High
Artificial Neural Network	Very High	Very High	Very High	Very High
Random Forest	Very High	Very High	Very High	Very High

Analysis

The Table 3 shows, Artificial Neural Networks, Random Forests, and Support Vector Machines consistently achieved the highest intrusion detection performance. These algorithms effectively modeled nonlinear attack behavior and complex relationships among network traffic features. Random Forest classifiers demonstrated strong robustness through ensemble learning, while Artificial Neural Networks achieved superior performance for anomaly detection in high-dimensional datasets.

Conclusion and Discussion

The present study investigated the application of Big Data technologies and intelligent machine learning techniques for enhancing intrusion detection in modern computer network security through a systematic review of research published between 2008 and 2018. The primary objective was to examine how distributed Big Data processing, intelligent classification algorithms, anomaly detection techniques, and cybersecurity analytics can be integrated to develop an efficient intrusion detection framework capable of processing massive volumes of network traffic while accurately identifying cyber threats. Based on the findings of the reviewed literature, this study proposed the Big Data-Based Intelligent Intrusion Detection System (BDI-IDS), which integrates distributed data acquisition, Big Data storage, data preprocessing, intelligent machine learning classification, anomaly detection, security decision support, and continuous performance monitoring within a unified cybersecurity architecture. One of the major conclusions of this study is that the rapid expansion of computer networks, cloud computing, mobile communication, and Internet-based services has dramatically increased the complexity of cybersecurity. During the 2008–2018 period, organizations experienced exponential growth in network traffic generated by enterprise systems, cloud infrastructures, mobile devices, and web-based applications. This enormous increase in data volume exceeded the processing capabilities of conventional intrusion detection systems, creating an urgent need for scalable security architectures capable of analyzing distributed network traffic efficiently. Big Data technologies emerged as an effective solution by enabling distributed storage, parallel computation, and large-scale

data analytics, thereby transforming the way cybersecurity systems process network information. The literature reviewed in this study consistently demonstrates that traditional signature-based intrusion detection systems remain effective for identifying previously known cyberattacks but exhibit significant limitations when detecting unknown or evolving threats. Signature-based approaches depend on predefined attack patterns stored in signature databases. Consequently, they cannot effectively recognize zero-day attacks, polymorphic malware, advanced persistent threats, insider attacks, or newly emerging malicious behaviors that differ from previously observed attack signatures. This limitation has become increasingly critical as cybercriminals continuously develop sophisticated attack techniques designed to evade conventional security mechanisms. The reviewed studies indicate that machine learning has fundamentally transformed intrusion detection by enabling intelligent systems to learn complex attack characteristics directly from historical network data. Algorithms such as Decision Trees, Support Vector Machines (SVMs), Artificial Neural Networks (ANNs), Random Forests, Naïve Bayes, and K-Nearest Neighbors (KNN) consistently demonstrated superior classification performance compared with conventional rule-based detection systems. Artificial Neural Networks effectively modeled nonlinear relationships among network traffic features, while Support Vector Machines achieved high classification accuracy for high-dimensional cybersecurity datasets. Random Forest algorithms further improved detection robustness by combining multiple decision trees into an ensemble learning architecture capable of reducing classification variance and improving prediction reliability.

References

1. Apache Hadoop. (2008). *Apache Hadoop*. Apache Software Foundation. https://hadoop.apache.org/
2. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials, 18*(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502
3. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM, 51*(1), 107–113. https://doi.org/10.1145/1327452.1327492
4. Javaid, A., Niyaz, Q., Sun, W., & Alam, M. (2016). A deep learning approach for network intrusion detection system. *Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (BICT 2016)*, 21–26. https://doi.org/10.4108/eai.3-12-2015.2262516
5. Liao, H.-J., Lin, C.-H. R., Lin, Y.-C., & Tung, K.-Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications, 36*(1), 16–24. https://doi.org/10.1016/j.jnca.2012.09.004
6. Ring, M., Wunderlich, S., Grüdl, D., Landes, D., & Hotho, A. (2018). A survey of network-based intrusion detection datasets. *Computers & Security, 86*, 147–167. https://doi.org/10.1016/j.cose.2019.06.005

7. Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A deep learning approach to network intrusion detection. *IEEE Transactions on Emerging Topics in Computational Intelligence, 2*(1), 41–50. https://doi.org/10.1109/TETCI.2017.2772792
8. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. *Proceedings of the IEEE Symposium on Security and Privacy*, 305–316. https://doi.org/10.1109/SP.2010.25
9. Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 dataset. *Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA 2009)*, 1–6. https://doi.org/10.1109/CISDA.2009.5356528
10. Tsai, C.-F., Hsu, Y.-F., Lin, C.-Y., & Lin, W.-Y. (2009). Intrusion detection by machine learning: A review. *Expert Systems with Applications, 36*(10), 11994–12000. https://doi.org/10.1016/j.eswa.2009.03.012
11. Wu, S. X., & Banzhaf, W. (2010). The use of computational intelligence in intrusion detection systems: A review. *Applied Soft Computing, 10*(1), 1–35. https://doi.org/10.1016/j.asoc.2009.06.019
12. Zuech, R., Khoshgoftaar, T. M., & Wald, R. (2015). Intrusion detection and Big Heterogeneous Data: A survey. *Journal of Big Data, 2*(3), 1–41. https://doi.org/10.1186/s40537-015-0013-4
13. Lee, W., Stolfo, S. J., & Mok, K. W. (1999). A data mining framework for building intrusion detection models. *Proceedings of the IEEE Symposium on Security and Privacy*, 120–132. https://doi.org/10.1109/SECPRI.1999.766909
14. Patcha, A., & Park, J.-M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks, 51*(12), 3448–3470. https://doi.org/10.1016/j.comnet.2007.02.001
15. Scarfone, K., & Mell, P. (2007). *Guide to Intrusion Detection and Prevention Systems (IDPS)* (NIST Special Publication 800-94). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-94