A Multi-level Features Fusion Model for Network Communication based on Machine Learning

Mahmoud A. Zaher[1,*], Nabil M. Eldakhly[2]

[1] Faculty of Artificial Intelligence, Egyptian Russian University (ERU), Cairo, Egypt

[2] Department of Computer and Information Systems, Sadat Academy for Management Sciences, Cairo, Egypt

Emails: Mahmoud.zaher@eru.edu.eg ; nmeldakhly@yahoo.com

Abstract

Today's societies couldn't function without elaborate networks of communication. Many problems remain unresolved, but novel approaches to these problems are constantly being offered. Many of the problems plaguing existing works, such as high characteristic design cost, challenging feature selection, poor real-time performance, etc., stem from their focus on a wide range of characteristics. Worse still, the difficulty in training models due to data imbalance results in a poor detection rate for aberrant samples. To achieve a more effective and robust model, we present a multi-level feature fusion (MFFusion) model that utilizes a combination of data temporal, byte, and statistical characteristics to extract relevant information from different angles. Too far, MFFusion has outperformed the state-of-the-art on several real-world network datasets in terms of prediction performance and false alarm rate. We also use MFFusion for anomaly detection in an IoT network, using the most recent IoT malicious traffic information. The experimental results demonstrate the adaptability of MFFusion and its suitability for identifying network anomalies in an IoT context with system performance.

Keywords: Network Communication; IoT; Multi-Level Fusion; Deep Learning; Machine Learning

1. Introduction

The Internet, 4G/5G cellular technology, and the rapidly expanding Internet of Things (IoT) are just a few examples of the ubiquity of communication networks in modern culture. The exponential development of communication systems has outpaced the wildest expectations of its creators. Nearly 2 of the world's population, for instance, will have access to the Internet by 2023, according to the Cisco Annual Latest Report (2018-2023) White Paper. Such massive networks would be very difficult to maintain and control, and the emergence of new kinds of networks just adds to the complexity of the situation[1]–[3]. For instance, manual setup often becomes impractical or wasteful in today's networks. While studies of communication networks have been conducted for quite some time, the field is still very much alive and kicking, with recent innovations including Software Defined (SDN) and Room Integrated Networks (SAGIN). Virtual network encapsulation in SDN is an example of a newer difficulty that might arise alongside more established ones like routing and traffic shaping, power control, and resource allocation[4]–[6].

Many new approaches, including deep learning, have been brought to the networking area to address these difficulties. In many applications, deep learning, as reflected by neural networks, has been shown to be very effective. This is notably true in the areas of image identification, language processing, and series data issues. Many types of communication networks employ deep learning models for a variety of purposes, including but not limited to network architecture, traffic prediction,

Doi : https://doi.org/10.54216/IJWAC.050103
Received: March 09, 2022 Accepted: October 03, 2022

allocating resources, etc. Most deep neural systems are created for data with a Euclidean structure, such as photos and videos, therefore they aren't completely using the network topological structure in this research. Recent years have seen the proposal of graph-based deep learning, in the form of Graph Neural Networks (GNNs), for non-Euclidean structural data in an effort to address this problem[7]–[9]. More recently, GNNs have been integrated with deep learning for decision-making across a range of domains; for instance, GNNs are used to analyze graph information and enhance distributed computing's inter-coflow scheduling capability. The most frequent issues, such as network modeling, routing, and traffic prediction, are presented in two or three different examples as shown in figure 1.

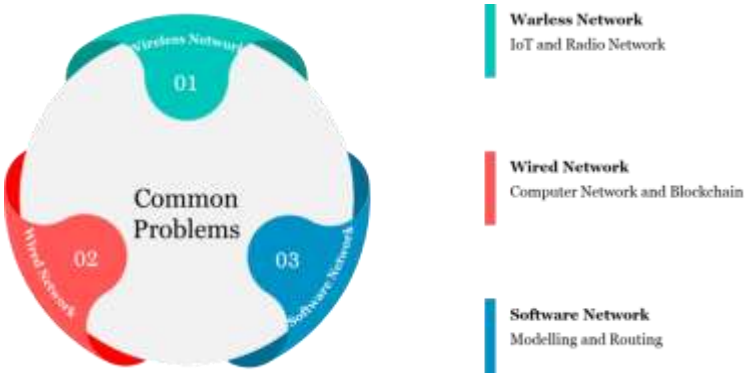

Figure 1: The common network communication problems.

1.1 Communication Network Traffic Detection

To ensure the safety of the internet, it is crucial to be able to identify malicious traffic and unusual patterns of network activity. From the viewpoint of the traffic carrier, anomalous network behavior is identical to that of any other network application, each of which is made up of a linear arrangement of network data packets forming a unique network flow[10]–[12].

Current network abnormality detection techniques face significant hurdles due to the ongoing development of network design and the exponential increase of network hardware. Many industries have adopted Internet of Things (IoT) solutions in recent years, and experts predict that IoT will be a game-changer in the next industrial revolution. On the other hand, a huge number of deployed IoT devices can only be protected to a limited degree. Most of these gadgets do not have any kind of built-in security system, so when they are connected to the Internet, they create thousands of potential entry points for hackers to exploit. As a result, concerns about network security have emerged as a major roadblock to the progress of future network infrastructure. However, the present logically centralized detection and prevention methodology and intrusion warning system for wireless sensor networks cannot satisfy the needs for stability, dissemination, resource limits, and low latency in the Internet of Things[13]–[16]. As the number of connected devices continues to rise, the amount of information that has to be stored, processed, and computed also increases, posing significant problems for centralized network aberration sensing devices. According to the study, by 2025 there will be 75.48 billion networked gadgets in use all over the world. By 2023, the number of gadgets linked to Internet protocol will exceed three times the number of people on the planet. It is expected that ultimately cloud data centers will not be able to purchase such a massive quantity of processing[17]–[19]. Edge computing and the Internet of Things in the network's periphery are two examples of the based on distributed cognitive computing network topologies offered by researchers as a solution to this issue. By using these tools, computation nodes may provide distributed, low-latency, and highly-available services to adjacent data sources. By moving computation to edge nodes, we can overcome IoT's resource limitations and meet its demands for computation, storage, and control. Because, like other services, centralized network anomaly-based architecture has a hard

Doi : https://doi.org/10.54216/IJWAC.050103
Received: March 09, 2022 Accepted: October 03, 2022

time adjusting to a decentralized setting, the best solution is to install "edge nodes," where harmful and abnormality detection is offloaded to safeguard devices directly linked to the network.

1.3 Quantum networks

Secure communication across nodes of varying quantum and conventional capabilities, as well as efficient distributed quantum computation, are made possible by quantum networks. To effectively transfer quantum information among communications nodes, a quantum wireless transmission network must be both a complicated system and able to provide a safe wireless network connection. Cheng et al. introduced the first quantum routing technique in a hierarchy routing protocol to transfer a quantum state from one node to another, even if the nodes did not share Bell pairings[20]–[22]. Yu et al. presented a routing system for a dispersed wifi router and a wireless ad hoc quantum telecommunication system. A quantum multi-hop telecommunication system based on entanglement switching and simultaneous measurement was recently reported by Wang et al. Quantum bridging using partly entangled states was suggested for use in hop-by-hop teleportation to cut down on the time investment[23]–[25]. Based on hybrid Werner states, Shi et al. suggested a quantum multi-hop network communication for arbitrary single qubits. A quantum protocol for multi-hop communication using partly entangled GHZ states was developed by Xiong et al. In order to facilitate quantum multi-hop communication, Zhan et al. suggested using W states and EPR couples. Two-qubit quantum multi-hop routing protocol using the compound GHZ-Bell channel was introduced by Zou et al. To facilitate one-qubit, two-qubit, and N-qubit atomic multi-hop telecommunication, Zhang et al. suggested a state that is a hybrid of an asymmetric W state and a Bell state. Researchers Yang et al. looked at using cluster states for mobile network communication and came up with a few different techniques for doing so using 1D, 2D, and 3D cluster states. In addition, mesh topology-based multi-hop classical translocation techniques have been developed[26], [27].

Despite the fact that the aforementioned schemes can be used to accomplish quantum multi-hop information exchange of a single or numerous quantum states by using different entangled states, they suffer from a major drawback: each pair of destination nodes on the path must share multi-qubit entangled states, such as a Bell country, Werner condition, partially intertwined GHZ state, W-Bell state, Quantum storage, preparations, and measurements should be minimized in their need for mobility in quantum nodes.

2. The Proposed Methodology

In this part, you will learn about the preprocessing techniques, model structure, and MFFusion's parameter settings that make it unique.

In the experiment section, three actual network datasets from ISCXIDS2012 are used. Labels, timeout, log data, as well as other statistical aspects of each packet are included in this dataset together with the original data collection files and explanation files (or flow). There is a substantial quantity of regular traffic interspersed with the anomalous samples, therefore the data preprocessing must simultaneously parse the PCAP file and the descriptive file before the label-matching operation can be performed. Figure 2 depicts the entire process flow.

Figure 2: The data pre-processing.

Doi : https://doi.org/10.54216/IJWAC.050103
Received: March 09, 2022 Accepted: October 03, 2022

Each dataset has a unique recording format for its accompanying explanation files.

For label matching to work, it must first read the description file, at which point it may use the date and quintuple data to find the relevant flow and label it appropriately.

After the segmentation process, each flow is treated as a separate example, which is a file that contains. Before feeding data into a model, it's crucial to settle on an appropriate vectorization approach. The model may be taught new abilities by using various vectorization strategies.

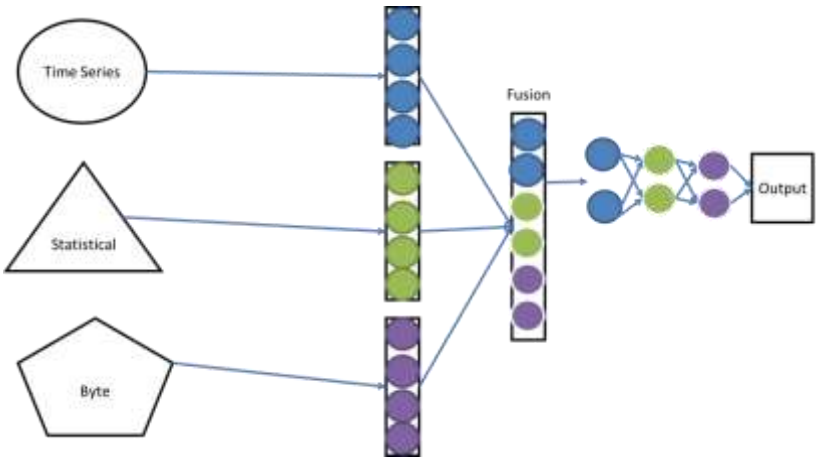

Figure 3: Multi-level fusion.

The vectorization section employs three distinct vectorization techniques to collect characteristics of varying granularity, allowing for the incorporation of viewpoint-specific data into the model. These vectors are taken straight from the raw data of the traffic, but in order to gain more useful abstract features, it is important to use deep learning techniques to craft unique model architectures for the purpose of automated feature extraction. Our research presents a multi-level features fusion model, the architecture of which is shown in Figure 3.

The comprehensiveness of findings is typically improved by analyzing data from many angles. In all, there are four sub-models that make up the MFFusion framework. Timing, byte, and statistical features are each taken into consideration by the first 3 network architectures. The characteristics acquired in the first three steps are fused into the fully-connected neural network in the fourth step, where they are used for joint learning. Table 1 shows the specific values used for the MFFusion model.

Table 1: Parameters of multi-level fusion

Network	Layer
Network	LSTM 1
	Dropout
	LSTM 2
	Dropout
	Linear
	Relu

Doi : https://doi.org/10.54216/IJWAC.050103
Received: March 09, 2022 Accepted: October 03, 2022

3. Results and discussion

Accuracy, recall, precision, and F1 are typical measures of performance in categorization tasks. However, in the realm of network malicious detection, the terms detection accuracy (DR) and rates of false alarms (FAR) are often employed. A reduced false alarm rate may minimize the cost of misjudgment, while a greater detection rate can avoid the omission of risky activities. Real-world applications of anomaly-based strive for a high detection rate of unusual samples with a low false alarm rate since aberrant attacks might have dire repercussions.

To learn all three properties simultaneously, the MFFusion model employs a trifecta of vectorization techniques and associated network topologies. The idea behind it is that broadening one's grasp of a topic may sometimes provide better results. This section discusses how to rate each feature's usefulness to the overall model.

The saliency map is used to get the model's gradient to the input in deep learning, a technique often used in computer vision to ascertain how each pixel in the input picture contributes to the model overall. Similarly, this section gets the total exact value of the gradients of each characteristic input to evaluate the importance of each grade feature in the system.

Figure 4 provides an alternative vantage point from which to examine the contributions of the three networks. The goal is to see how the empirical measures change when neural units are removed from one of the multiple network outputs (i.e., inputs to Transmit the signal) in a trained model.

Figure 4: Degree of performance level.

Figure 4 shows that eliminating the leverages of any one of the three Fusion networks would decrease DR and raise FAR, with the greatest effect shown in the Binary, Series Data network. It's consistent with what we found using saliency maps. Surprisingly, when the Statistical component is off, the ACC really improves. While its FAR increases from 0.43% to 1.80%, its DR decreases from 98.75%, making it less than ideal for detecting anomalies in a network. The impact of multicollinearity on DR and FAR is exacerbated since the ACC increase is the result of more problematic samples being misdiagnosed as normal. Because the system is more likely to glean data from the well-planned features extracted if there are insufficient examples for extracting information from deep learning features, removing the Statistics modules has a stronger influence on those aberrant kinds with fewer data.

Doi : https://doi.org/10.54216/IJWAC.050103
Received: March 09, 2022 Accepted: October 03, 2022

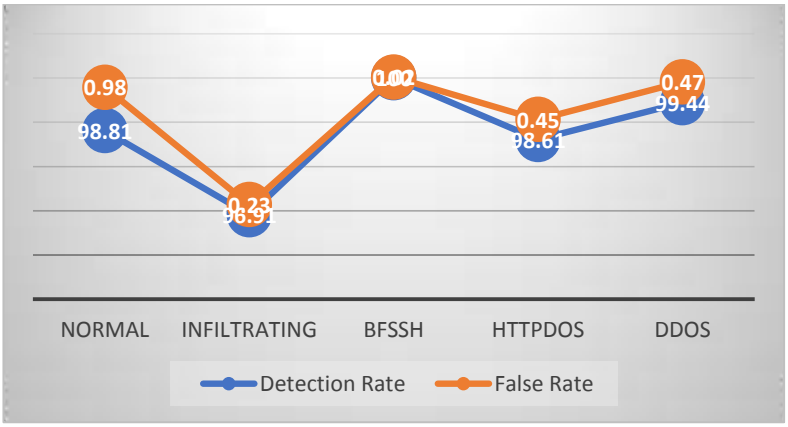

Figure 5: Overall Performance.

Figures 5, 6, and 7 provide comprehensive data from the experiments. A small drop in performance on ISCXIDS2012 did not prevent MFFusion from achieving a DR of 98.75% and a FAR of 0.43 percent.

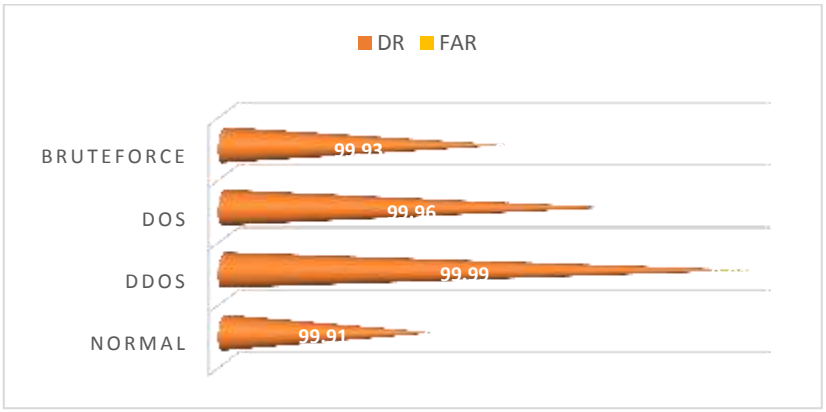

Figure 6: The second Performance

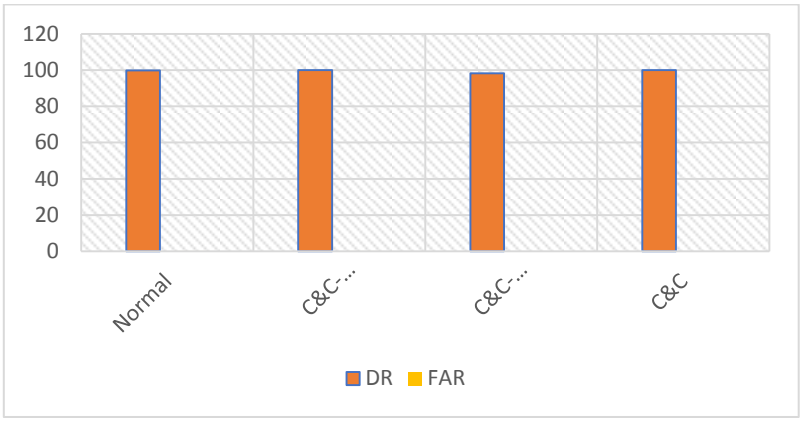

Figure 7: The third performance.

Anomaly detection in IoT networks was performed using MFFusion. Although the behavioral features and traffic distribution are shifting in the IoT scenario, overall performance met the processing capabilities, and certain categories even surpassed 100%.

This demonstrates MFFusion's adaptability and demonstrates that it can be used to identify anomalies in network traffic in an Internet of Things set.

Doi : https://doi.org/10.54216/IJWAC.050103
Received: March 09, 2022 Accepted: October 03, 2022

The contemporary edge IoT ecosystem is characterized by a plethora of IoT devices being linked to the Internet via edge gates, which provides a wealth of benefits to society as a whole. But malicious actors may also launch attacks on IoT devices when they are connected to the Internet. In addition, a significant number of networking devices would create hundreds of millions of messages per day, rendering the intrusion detection systems housed in the cloud services center inadequate to the requirements of the distributed manner. The MFFusion framework is suitable for deployment at the edge nodes. It will safeguard IoT devices that are hardwired into edge nodes by detecting aberrant network activity via data gathering and analysis. It diverts network attack monitoring from a central location to distributed nodes at the network's periphery. This means that jobs that need lots of resources like computation, storage, and administration may be moved to the network's periphery.

A detection system in IoT networks is a good fit for MFFusion. The effectiveness of MFFusion may also reach the processing capabilities in an IoT context, as shown by tests and investigations done on actual IoT harmful network datasets.

6. Conclusion

To achieve the rapid recognition of aberrant network traffic, this research provides a malicious traffic approach to detect based on MFFusion. The detection rate, the false alarm rate, and the resilience of MFFusion are all improved by the combination of three tiers of features: time, bytes, and short-term statistical characteristics. Having trouble collecting and correctly categorizing anomalous network samples is a common result of working in a network, which might cause a severe shortage of such samples. Finally, MFFusion may be implemented at edge nodes to offer network security assistance for things attached directly to the node, making it ideal for detecting malicious traffic in an IoT setting.

Future research is the integration of GNNs with other forms of AI. Combinations of GNN and GRU for modeling, as well as GNN and DRL for allocating resources, routing, and VNE, may be shown in this overview. While GNNs do have certain benefits, such as the capacity to understand topological relationships and the ability to generalize for unknown network typologies, they are not a silver bullet. For scenarios in which there is a dearth of training examples or where it would be prohibitively costly to obtain actual data, GANs provide a potential answer. While GANs have seen extensive application in other domains, such as image and video analysis, our review does not find any examples of their deployment in communications infrastructure using a mix of GANs and GNNs. The Automatic Deep - learning (AutoML) method is another case in point; it may be used to automatically optimize the GNN's settings.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] E. W. L. Cheng, H. Li, P. E. D. Love, and Z. Irani, “Network communication in the construction industry,” *Corporate Communications: An International Journal*, vol. 6, no. 2, pp. 61–70, 2001.

[2] M. Schlichting, M. M. Díaz, J. Xin, and M. Rosbash, “Neuron-specific knockouts indicate the importance of network communication to Drosophila rhythmicity,” *Elife*, vol. 8, p. e48301, 2019.

[3] M. Kang, G. Yang, Y. Yoo, and C. Yoo, “TensorExpress: In-network communication scheduling for distributed deep learning,” in *2020 IEEE 13th international conference on cloud computing (CLOUD)*, 2020, pp. 25–27.

[4] X.-B. Chen, Y.-L. Wang, G. Xu, and Y.-X. Yang, “Quantum network communication with a novel discrete-time quantum walk,” *Ieee Access*, vol. 7, pp. 13634–13642, 2019.

[5] S. K. Tripathi, M. Kumar, and A. Kumar, "Graphene-based tunable and wideband terahertz antenna for wireless network communication," *Wireless Networks*, vol. 25, no. 7, pp. 4371–4381, 2019.

[6] W. Li, Z. Xie, J. Zhao, and P. K. Wong, "Velocity-based robust fault-tolerant automatic steering control of autonomous ground vehicles via adaptive event-triggered network communication," *Mechanical Systems and Signal Processing*, vol. 143, p. 106798, 2020.

[7] Y.-C. Sun and G.-H. Yang, “Event-triggered state estimation for networked control systems with lossy network communication,” *Information Sciences*, vol. 492, pp. 1–12, 2019.

[8] V. N. Poole, O.-Y. Lo, T. Wooten, I. Iloputaife, L. A. Lipsitz, and M. Esterman, “Motor-cognitive neural network communication underlies walking speed in community-dwelling older adults,” *Frontiers in aging neuroscience*, vol. 11, p. 159, 2019.

Doi : https://doi.org/10.54216/IJWAC.050103
Received: March 09, 2022 Accepted: October 03, 2022

[9] H. Lin, Z. T. Kalbarczyk, and R. K. Iyer, "Raincoat: Randomization of network communication in power grid cyberinfrastructure to mislead attackers," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 4893–4906, 2018.
[10] T. Kustermann, T. Popov, G. A. Miller, and B. Rockstroh, "Verbal working memory-related neural network communication in schizophrenia," *Psychophysiology*, vol. 55, no. 9, p. e13088, 2018.
[11] H. Huang *et al.*, "Deep learning for physical-layer 5G wireless techniques: Opportunities, challenges, and solutions," *IEEE Wireless Communications*, vol. 27, no. 1, pp. 214–222, 2019.
[12] H. Wu, X. Li, and Y. Deng, "Deep learning-driven wireless communication for edge-cloud computing: opportunities and challenges," *Journal of Cloud Computing*, vol. 9, no. 1, pp. 1–14, 2020.
[13] Z. Tao and Q. Li, "{eSGD}: Communication Efficient Distributed Deep Learning on the Edge," 2018.
[14] G. Thamilarasu and S. Chawla, "Towards deep-learning-driven intrusion detection for the internet of things," *Sensors*, vol. 19, no. 9, p. 1977, 2019.
[15] Z. Lv, A. K. Singh, and J. Li, "Deep learning for security problems in 5G heterogeneous networks," *IEEE Network*, vol. 35, no. 2, pp. 67–73, 2021.
[16] A. Koloskova, T. Lin, S. U. Stich, and M. Jaggi, "Decentralized deep learning with arbitrary communication compression," *arXiv preprint arXiv:1907.09356*, 2019.
[17] J. Chen, K. Li, Q. Deng, K. Li, and S. Y. Philip, "Distributed deep learning model for intelligent video surveillance systems with edge computing," *IEEE Transactions on Industrial Informatics*, 2019.
[18] M. Mahdavisharif, S. Jamali, and R. Fotohi, "Big data-aware intrusion detection system in communication networks: a deep learning approach," *Journal of Grid Computing*, vol. 19, no. 4, pp. 1–28, 2021.
[19] N. Shazeer *et al.*, "Mesh-TensorFlow: Deep learning for supercomputers," *Advances in neural information processing systems*, vol. 31, 2018.
[20] H. Geng, H. Liu, L. Ma, and X. Yi, "Multi-sensor filtering fusion meets censored measurements under a constrained network environment: advances, challenges, and prospects," *International Journal of Systems Science*, vol. 52, no. 16, pp. 3410–3436, 2021.
[21] C. Yuan, W. Yue-ping, T. Xue-mei, and Z. Xiao-fang, "An evaluation method for network communication system efficiency based on multi-source information fusion," in *2018 37th Chinese Control Conference (CCC)*, 2018, pp. 4305–4309.
[22] T. Lin, P. Wu, and F. Gao, "Information security of flowmeter communication network based on multi-sensor data fusion," *Energy Reports*, vol. 8, pp. 12643–12652, 2022.
[23] A.-M. Yang, X.-L. Yang, J.-C. Chang, B. Bai, F.-B. Kong, and Q.-B. Ran, "Research on a fusion scheme of the cellular network and wireless sensor for cyber-physical social systems," *Ieee Access*, vol. 6, pp. 18786–18794, 2018.
[24] T. Li, J. M. Corchado, and S. Sun, "Partial consensus and conservative fusion of Gaussian mixtures for distributed Ph.D. fusion," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 5, pp. 2150–2163, 2018.
[25] Y. Yuan and Y. Li, "Research on Spatial Agglomeration Characteristics of Aerospace Cultural and Creative Industries in Smart City under Multidata Fusion," *Security and Communication Networks*, vol. 2022, 2022.
[26] R. Xu, H. Xiang, X. Xia, X. Han, J. Li, and J. Ma, "Opv2v: An open benchmark dataset and fusion pipeline for perception with vehicle-to-vehicle communication," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 2583–2589.
[27] G. Li, Z. Yan, Y. Fu, and H. Chen, "Data fusion for network intrusion detection: a review," *Security and Communication Networks*, vol. 2018, 2018.

Doi : https://doi.org/10.54216/IJWAC.050103
Received: March 09, 2022 Accepted: October 03, 2022