RESEARCH ARTICLE OPEN ACCESS

Malicious Attack Detection by Convolutional Deep Learning Model for Web Applications

Madhankumar Y
Government Arts College (Autonomus)
Kumbakonam-Thanjavur

ABSTRACT
Wireless Sensor Network (WSN) is profoundly imperative for securing network protection. Profoundly basic attacks of different types have been reported in wireless sensor network till now by numerous researchers. The Distributed Denial of Service (DDOS) attack is a standard type of attack in WSN, comprising two composes, to be specific; passive attacks and active attacks, of which the last can cause more noteworthy risk. General goal of this paper to recognize the DDoS attack in WSN utilizing creative clustering and threshold-based detection process with consider the routing protocol as LEACH. At first network is isolated into number of clusters, each cluster has one header and the header is specifically conveyed to destination. From that the nodes which often make trouble and in light of their miss ratio they will be eliminated from the network. The viability of our proposed detection framework is assessed utilizing some performance measures and it's executed in network simulator.

Keywords:

I. INTRODUCTION

A Wireless Sensor Network (WSN) is a wireless network comprising of spatially conveyed self-ruling devices utilizing sensors to screen physical or ecological conditions [1]. LEACH protocol is the primary convention of various levelled routing which proposed data combination; it is of point of reference hugeness in clustering routing protocol. Routing procedures and security issues are awesome research challenge [2]. A network that passage data to another network, that is it get the information from one network imitate it into another network through passage that specific network may befuddled because of this activity. Around then programmer may effectively enter and do abuse inside the network [3, 4]. A DDOS attack (otherwise called a Distributed Denial of Service attack) is a sort of web attack that tries to disturb the typical capacity of the focused-on PC network. Albeit a few arrangements were anticipated to disentangle the DDoS drawback and a couple of sorts of the attack's region unit so [5,6] countered, DDoS attacks still be principle danger inside the net. The Denial of service is the active type of attack in which malicious node flood the legitimate nodes with the rough packets to reduce network performance [7].

The distributed denial of service is the advance type of DOS attack [7] in which malicious node choose its slave and slaves will flood the legitimate node which the rough packets and it reduce network performance [7-10]. When all the nodes start transmitting data in the network, and when the DDOS attack is triggered in the network and throughput of the network get reduced to threshold value then malicious node detection process starts. In the process of malicious node detection, the nodes which are sending data above the threshold [11] value are considered as malicious node and technique of watch dog is applied that whether these nodes are sending data packets or control packets [12].

Feature selection and dimensionality reduction are two method that aim at solving these problems by reducing the number of features and thus the dimensionality of the data [13]. The most common and useful unsupervised feature transformation is PCA
proposed by Pearson in the early 20th century. In this context, users' behaviours, including both normal and malicious ones, are defined and represented using the accumulated packets throughout a network [14-20]. We believe that a combination of traditional security systems with machine learning techniques can provide a new intrusion protection system that is meaningfully and mathematically justified and provides a secure platform for all users [21,22].The major challenges in the field of malicious network activity detection are a huge volume of network traffic, diversity due to new attacks, and reduced performance of low-frequency attacks due to a high imbalance between various classes of attacks [23]. Deep learning techniques can help to tackle these challenges as they have the ability to model complex relationships and concepts using multiple levels of representation [24-30].

II. LITERATURE REVIEW

In 2017 NejlaRouissi et al. [31] have proposed a novel energy proficient and data integrity rendition of LEACH construct routing protocol with respect to Watermarking for wireless sensor networks since LEACH routing protocol does not think about the security viewpoint and the protected enhanced LEACH works were construct just in light of cryptographic systems. The hybrid proposed approach in view of the Watermarking-LEACH accomplishes data respectability as well as Energy-Efficient. It was the main pattern that endeavours to include security-based watermarking to LEACH routing protocol.

In 2017 Amar Meryem et al. [32] the fundamental objective of this work is the recognizable proof and forecast of assaults and vindictive practices by breaking down, ordering and marking recorded exercises in log documents. This paper utilizes MapReduce programming to earlier every client conduct, it additionally utilizes K-Means algorithm to bunch obscure occasions and K-NN regulated learning on NSLKDD database to characterize unlabeled classes. This procedure order identified assaults amid the training stage and along these lines enables us to know, with a specific likelihood, if an obscure conduct is related to the presence of an attack.
Feature Selection in Intrusion Detection Grey Wolf Optimizer by E M Roopa Devi; R C Suganthe (2017) [33]. The creator has utilized gray wolf optimizer a swarm-based optimization technique to look through the element space to discover ideal component subset that enhances classification accuracy. At to start with, the grey wolf streamlining agent utilizes channel-based standards to discover arrangements with minor excess that are depicted by shared data. At the later stage optimization wrapper approach is utilized for controlling classifier execution. The execution of dim wolf enhancer is measured and contrasted against a few other metaheuristic algorithms and the assistance of NSL KDD Dataset [34].

Intrusion Detection is a security technique, used to monitor and analysed network traffic in order to detect network violation by Rachana Sharma et al (2016) [19]. Machine learning methods are to detect intrusion which can scale up to assemble such frameworks. There are numerous algorithms one can decide on relying on the requirements of the framework. This paper manages Naïve Bayes and K-Nearest Neighbour classifier in MapReduce structure and their execution contrasted and WEKA usage.
In 2019 Mahdi Rabbani et al.[35] have been proposed a new approach to improve the capability of Cloud service providers to model users' behaviours. We applied a particle swarm optimization-based probabilistic neural network (PSO-PNN) for the detection and recognition process. In the first module of the recognition process, we meaningfully converted the users' behaviours to an understandable format and then classified and recognized the malicious behaviours by using a multi-layer neural network. We took advantage of the UNSW-NB15 dataset to validate the proposed solution by characterizing different types of malicious behaviours exhibited by users.
The detection of malicious code is becoming increasingly crucial, and current methods of detection require much improvement by Cui, Z. et al.in 2029 [36]. This paper proposes a method to advance the detection of malicious code using convolutional neural networks (CNNs) and intelligence algorithm. The CNNs are used to identify and classify grayscale images converted from executable files of malicious code. Non-dominated Sorting Genetic Algorithm II (NSGA-II) is then employed to deal with the data imbalance of malware families. A series of experiments are designed for malware image data from Vision Research Lab.

III. RESEARCH GAP

One of the real reasons that make the DDoS attacks across the board and simple in the Internet is the accessibility of attacking tools and the effectiveness of these tools to create attacking traffic. There are a wide range of DDoS attack tools on the Internet that enable attackers to execute attacks on the objective framework [37,38]. The http flood attack causes large number of resources to be allotted in response to the few requests from the clients. Placing check on the allocation process reduced the problem but also reduces the resource consumption [40]. From the existing literatures many machine learning and supervised classifiers are used to detect the attack in network system, because the detection level is low, to overcome this we are proposed new methodology.

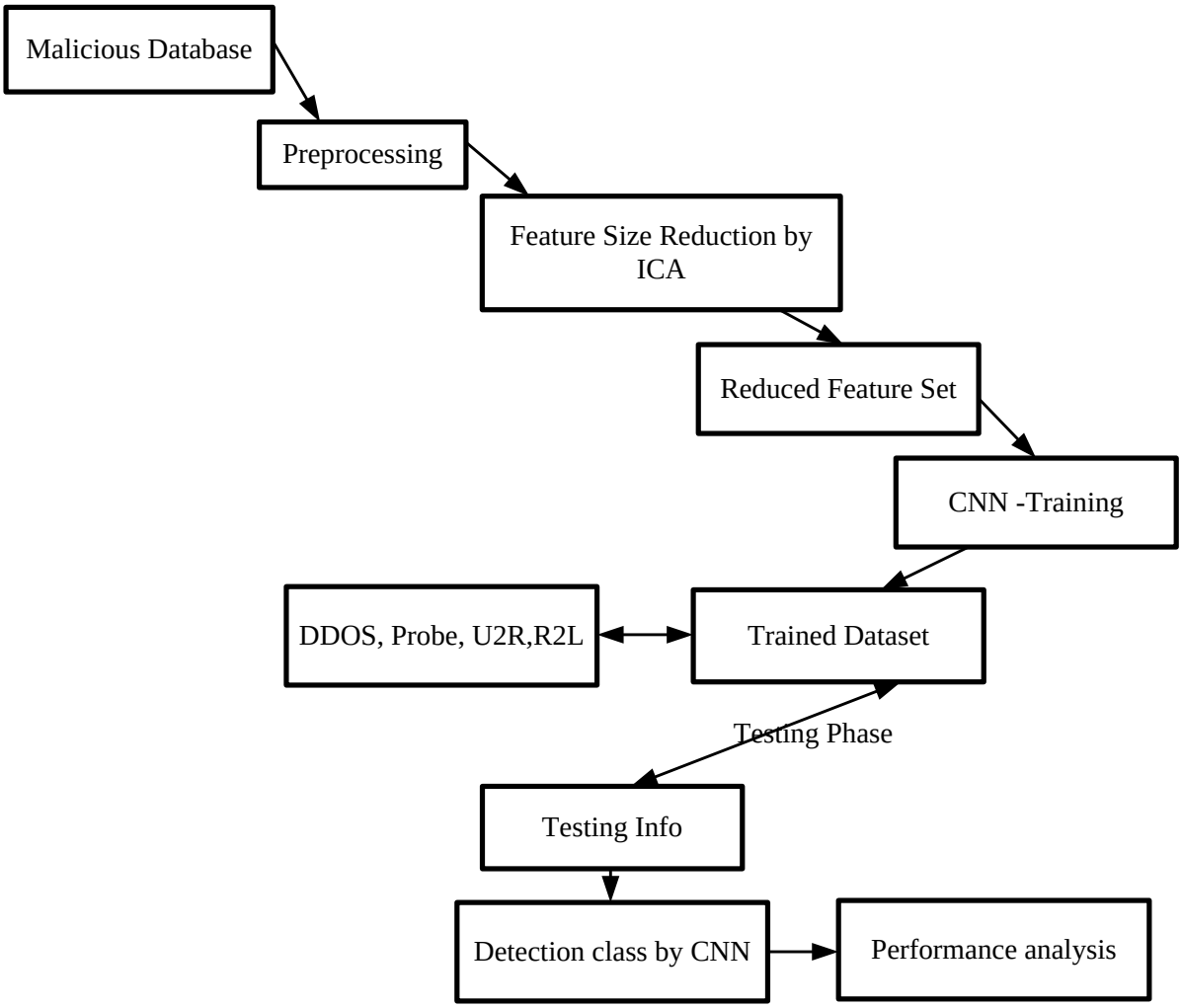
IV. ATTACK DETECTION METHODOLOGY

A passive attacker tunes in to the channel and may get to the packet containing mystery data exchanging from source to goals. The virtual topology network is industrialized that envelops an arrangement of different sensor nodes, one base station and server considered, this node are gathered in view of various portable network utilizing Machine learning and deep learning models [41,42]. This research work we are considered the KDD cup attack detection database are considered for attack detection and prevention modelling. For this proposed research model detecting the malicious data, initially perform the pre-processing to remove the unwanted data for detection stage [43]. In second phase feature reduction technique used to select the features by Incremental component analysis (ICA) technique, this approach based on the mutual information strategy for reduce the feature subsets [44-50]. The goal of feature subspace projections is to improve classifier robustness by reducing data dimensionality in order to facilitate better generalization, as well as reducing the learning and operating complexity of the classifiers. While doing so, classification performance must not be compromised by throwing away components that provide useful information regarding the class labels. At finally reduced feature based Convolutional Neural Network (CNN) to detect the DDOS and some attacks from KDD set [51-55].

This CNN structure comprises of three-layer capacities which are a convolutional layer, pooling layer lastly completely connected layers. The convolutional layers fill in as feature extractors, and subsequently, they take in the feature portrayals of their input images. The neurons in the convolutional layers are conceived into feature maps. Every neuron in a feature delineates a responsive field, which is associated with an area of neurons in the past layer by means of an arrangement of trainable weights [56]. Generally, detection process having Training and testing phases, this method is derived from cross-validation, with a subset of the available data kept out and used for testing on N number of folds. The output

classes are DDOS, Probe, U2R, R2L from KDD database, this model evaluated by confusion matrix information like TP, TN, FP and FN and also proposed research work compared with random forest, Neural Network, Naive bays and some other detection approaches [57].

Proposed Block Diagram



V. DATABASE DESCRIPTION

The data captured in DARPA'98 IDS evaluation program. DARPA'98 is about 4 gigabytes of compressed raw (binary) tcpdump data of 7 weeks of network traffic, which can be processed into about 5 million connection records, each with about 100 bytes. The two weeks of test data have around 2 million connection records. KDD training dataset consists of approximately 4,900,000 single connection vectors each of which contains 41 features and is labeled as either normal or an attack, with exactly one specific attack type.

Features for Attack detection

- To ncapsulates all the attributes that can be extracted from a TCP/IP connection. Most of these features leading to an implicit delay in detection.
- Traffic features category includes features that are computed with respect to a window interval and is divided into two groups: "same host" features: examine only the connections in the past 2 seconds that have the same destination host as the current connection, and calculate statistics related to protocol behavior, service, etc."same service" features: examine only the connections in the past 2 seconds that have the same service as the current connection [58].
- DoS and Probing attacks, the R2L and U2R attacks don't have any intrusion frequent sequential patterns. This is because the DoS and Probing attacks involve many connections to some host(s) in a very short period of time; however the R2L and U2R attacks

are embedded in the data portions of the packets, and normally involves only a single connection [59,60].

REFERENCE

[1] Chen, Y. T., Chen, C. H., Wu, S., & Lo, C. C. (2019). A two-step approach for classifying music genre on the strength of AHP weighted musical features. Mathematics, 7(1), 19.

[2] Elhoseny, M., Shankar, K., &Uthayakumar, J. (2019). Intelligent diagnostic prediction and classification system for chronic kidney disease. Scientific reports, 9(1), 1-14.

[3] Sivakumar, P., Velmurugan, S. P., & Sampson, J. (2020). Implementation of differential evolution algorithm to perform image fusion for identifying brain tumor.

[4] Khamparia, A., Gupta, D., Nguyen, N. G., Khanna, A., Pandey, B., & Tiwari, P. (2019). Sound classification using convolutional neural network and tensor deep stacking network. IEEE Access, 7, 7717-7727.

[5] Jansirani, A., Rajesh, R., Balasubramanian, R., & Eswaran, P. (2011). Hi-tech authentication for pslette images using digital signature and data hiding. Int. Arab J. Inf. Technol., 8(2), 117-123.

[6] Jain, R., Gupta, D., & Khanna, A. (2019). Usability feature optimization using MWOA. In International conference on innovative computing and communications (pp. 453-462). Springer, Singapore.

[7] Shankar, K., & Lakshmanaprabu, S. K. (2018). Optimal key based homomorphic encryption for color image security aid of ant lion optimization algorithm. International Journal of Engineering & Technology, 7(9), 22-27.

[8] Lyu, L., & Chen, C. H. (2020, July). Differentially Private Knowledge Distillation for Mobile Analytics. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 1809-1812).

[9] Poonkuntran, S., Rajesh, R. S., & Eswaran, P. (2011). Analysis of difference expanding method for medical image watermarking. In International Symposium on Computing, Communication, and Control (ISCCC 2009) (Vol. 1, pp. 31-34).

[10] Sampson, J., & Velmurugan, S. P. (2020, March). Analysis of GAA SNTFT with Different Dielectric Materials. In 2020 5th International Conference on Devices, Circuits and Systems (ICDCS) (pp. 283-285). IEEE.

[11] Elhoseny, M., Bian, G. B., Lakshmanaprabu, S. K., Shankar, K., Singh, A. K., & Wu, W. (2019). Effective features to classify ovarian cancer data in internet of medical things. Computer Networks, 159, 147-156.

[12] Gochhayat, S. P., Kaliyar, P., Conti, M., Tiwari, P., Prasath, V. B. S., Gupta, D., & Khanna, A. (2019). LISA: Lightweight context-aware IoT service architecture. Journal of cleaner production, 212, 1345-1356.

[13] Dutta, A. K., Elhoseny, M., Dahiya, V., & Shankar, K. (2020). An efficient hierarchical clustering protocol for multihop Internet of vehicles communication. Transactions on Emerging Telecommunications Technologies, 31(5), e3690.

[14] Anand Nayyar, Vikram Puri, Nhu Gia Nguyen, BioSenHealth 1.0: A Novel Internet of Medical Things (IoMT) Based Patient Health Monitoring System, Lecture Notes in Networks and Systems. Springer, 2019

[15] Shankar, K., Lakshmanaprabu, S. K., Khanna, A., Tanwar, S., Rodrigues, J. J., & Roy, N. R. (2019). Alzheimer detection using Group Grey Wolf Optimization based features with convolutional classifier. Computers & Electrical Engineering, 77, 230-243.

[16] Paramathma, M. K., Pravin, A. C., Rajarajan, R., & Velmurugan, S. P. (2019, April). Development and Implementation of Efficient Water and Energy Management System for Indian Villages. In 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS) (pp. 1-4). IEEE.

[17] Chen, C. H., Song, F., Hwang, F. J., & Wu, L. (2020). A probability density function generator based on neural networks. Physica A: Statistical Mechanics and its Applications, 541, 123344.

[18] Kathiresan, S., Sait, A. R. W., Gupta, D., Lakshmanaprabu, S. K., Khanna, A., & Pandey, H. M. (2020). Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. Pattern Recognition Letters.

[19] Gupta, D., & Ahlawat, A. K. (2016). Usability determination using multistage fuzzy system. Procedia Comput Sci, 78, 263-270.

[20] Amira S. Ashour, Samsad Beagum, Nilanjan Dey, Ahmed S. Ashour, Dimitra Sifaki Pistolla, Gia Nhu Nguyen, Dac-Nhuong Le, Fuqian Shi (2018), Light Microscopy Image De-noising using Optimized LPA-ICI Filter, Neural Computing and Applications, Vol.29(12), pp 1517–1533, Springer, ISSN: 0941-0643. (SCIE IF 4.664, Q1)

[21] Pan, M., Liu, Y., Cao, J., Li, Y., Li, C., & Chen, C. H. (2020). Visual Recognition Based on Deep Learning for Navigation Mark Classification. IEEE Access, 8, 32767-32775.

[22] Chen, C. H., Hwang, F. J., & Kung, H. Y. (2019). Travel time prediction system based on data

clustering for waste collection vehicles. IEICE TRANSACTIONS on Information and Systems, 102(7), 1374-1383.

[23] Shankar, K., & Eswaran, P. (2015). ECC based image encryption scheme with aid of optimization technique using differential evolution algorithm. Int J Appl Eng Res, 10(55), 1841-5.

[24] Anand Nayyar, Vikram Puri, Nhu Gia Nguyen, Dac Nhuong Le, Smart Surveillance Robot for the Real Time Monitoring and Control System in Environment and Industrial Applications, Advances in Intelligent System and Computing, pp 229-243, Springer

[25] Le Nguyen Bao, Dac-Nhuong Le, Gia Nhu Nguyen, Vikrant Bhateja, Suresh Chandra Satapathy (2017), Optimizing Feature Selection in Video-based Recognition using Max-Min Ant System for the Online Video Contextual Advertisement User-Oriented System, Journal of Computational Science, Elsevier ISSN: 1877-7503. Vol.21, pp.361-370. (SCIE IF 2.502, Q1)

[26] Chakchai So-In, Tri Gia Nguyen, Gia Nhu Nguyen: Barrier Coverage Deployment Algorithms for Mobile Sensor Networks. Journal of Internet Technology 12/2017; 18(7):1689-1699.

[27] Le, D.-N.a, Kumar, R.b, Nguyen, G.N., Chatterjee, J.M.d, Cloud Computing and Virtualization, DOI: 10.1002/9781119488149, Wiley.

[28] Bhateja, V., Gautam, A., Tiwari, A., Nhu, N.G., Le, D.-N, Haralick features-based classification of mammograms using SVM, Advances in Intelligent Systems and Computing, Volume 672, 2018, Pages 787-795.

[29] Khamparia, A., Saini, G., Gupta, D., Khanna, A., Tiwari, S., & de Albuquerque, V. H. C. (2020). Seasonal crops disease prediction and classification using deep convolutional encoder network. Circuits, Systems, and Signal Processing, 39(2), 818-836.

[30] Uthayakumar, J., Elhoseny, M., & Shankar, K. (2020). Highly Reliable and Low-Complexity Image Compression Scheme Using Neighborhood Correlation Sequence Algorithm in WSN. IEEE Transactions on Reliability.

[31] Huyen, D.T.T., Binh, N.T., Tuan, T.M., Nguyen, G.N, Dey, N., Son, L.H, Analyzing trends in hospital-cost payments of patients using ARIMA and GIS: Case study at the Hanoi Medical University Hospital, Vietnam, Journal of Medical Imaging and Health Informatics, 7(2), pp. 421-429.

[32] Van, V.N., Chi, L.M., Long, N.Q., Nguyen, G.N., Le, D.-N, A performance analysis of openstack open-source solution for IaaS cloud computing, Advances in Intelligent Systems and Computing, 380, pp. 141-150.

[33] Shankar, K., & Eswaran, P. (2016, January). A new k out of n secret image sharing scheme in visual cryptography. In 2016 10th International Conference on Intelligent Systems and Control (ISCO) (pp. 1-6). IEEE.

[34] Dey, N., Ashour, A.S., Chakraborty, S., Le, D.-N., Nguyen, G.N, Healthy and unhealthy rat hippocampus cells classification: A neural based automated system for Alzheimer disease classification, Journal of Advanced Microscopy Research, 11(1), pp. 1-10

[35] Velmurugan, S. P., & Rajasekaran, P. S. M. P. (2017). CLASSIFICATION OF BRAIN TUMOR USING MULTIMODAL FUSED IMAGES AND PNN. International Journal of Pure and Applied Mathematics, 115(6), 447-457.

[36] Shankar, K., Elhoseny, M., Perumal, E., Ilayaraja, M., & Kumar, K. S. (2019). An Efficient Image Encryption Scheme Based on Signcryption Technique with Adaptive Elephant Herding Optimization. In Cybersecurity and Secure Information Systems (pp. 31-42). Springer, Cham.

[37] Wu, L., Chen, C. H., & Zhang, Q. (2019). A mobile positioning method based on deep learning techniques. Electronics, 8(1), 59.

[38] Rouissi, N. and Gharsellaoui, H., 2017. Improved hybrid LEACH based approach for preserving secured integrity in wireless sensor networks. Procedia computer science, 112, pp.1429-1438.

[39] Meryem, A., Samira, D. and Mouad, L., 2017. A novel approach in detecting intrusions using NSLKDD database and MapReduce programming. Procedia Computer Science, 110, pp.230-235.

[40] E M Roopa Devi and R C Suganthe,"Feature Selection in Intrusion Detection Grey Wolf Optimizer",Asian Journal of Research in Social Sciences and Humanities,Vol.7,No.3,pp.671-682,2017.

[41] Sharma, Rachana, et al. "Towards MapReduce based classification approaches for Intrusion Detection." Cloud System and Big Data Engineering (Confluence), 2016 6th International Conference. IEEE, 2016.

[42] Rabbani, M., Wang, Y. L., Khoshkangini, R., Jelodar, H., Zhao, R., & Hu, P. (2019). A hybrid machine learning approach for malicious behaviour detection and recognition in cloud computing. Journal of Network and Computer Applications,

[43] Cui, Z., Du, L., Wang, P., Cai, X., & Zhang, W. (2019). Malicious code detection based on CNNs and multi-objective algorithm. Journal of Parallel and Distributed Computing.

[44] Lydia, E. L., Kumar, P. K., Shankar, K., Lakshmanaprabu, S. K., Vidhyavathi, R. M., & Maseleno, A. (2020). Charismatic document

clustering through novel K-Means non-negative matrix factorization (KNMF) algorithm using key phrase extraction. International Journal of Parallel Programming, 48(3), 496-514.

[45] Sujitha, B., Parvathy, V. S., Lydia, E. L., Rani, P., Polkowski, Z., & Shankar, K. (2020). Optimal deep learning based image compression technique for data transmission on industrial Internet of things applications. Transactions on Emerging Telecommunications Technologies, e3976.

[46] Lo, C. L., Chen, C. H., Hu, J. L., Lo, K. R., & Cho, H. J. (2019). A fuel-efficient route plan method based on game theory. Journal of Internet Technology, 20(3), 925-932.

[47] Kung, H. Y., Chen, C. H., Lin, M. H., & Wu, T. Y. (2019). Design of Seamless Handoff Control Based on Vehicular Streaming Communications. Journal of Internet Technology, 20(7), 2083-2097.

[48] Elhoseny, M., & Shankar, K. (2019). Reliable data transmission model for mobile ad hoc network using signcryption technique. IEEE Transactions on Reliability.

[49] Shanmugam, P., Rajesh, R. S., & Perumal, E. (2008, May). A reversible watermarking with low warping: an application to digital fundus image. In 2008 International Conference on Computer and Communication Engineering (pp. 472-477). IEEE.

[50] Shankar, K., & Elhoseny, M. (2019). Trust Based Cluster Head Election of Secure Message Transmission in MANET Using Multi Secure Protocol with TDES. J. UCS, 25(10), 1221-1239.

[51] Parvathy, V. S., Pothiraj, S., & Sampson, J. (2020). Optimal Deep Neural Network model based multimodality fused medical image classification. Physical Communication, 101119.

[52] Subbiah Parvathy, V., Pothiraj, S., & Sampson, J. (2020). A novel approach in multimodality medical image fusion using optimal shearlet and deep learning. International Journal of Imaging Systems and Technology.

[53] Parvathy, V. S., & Pothiraj, S. (2019). Multi-modality medical image fusion using hybridization of binary crow search optimization. Health Care Management Science, 1-9.

[54] Velmurugan, S. P., Sivakumar, P., & Rajasekaran, M. P. (2018). Multimodality image fusion using centre-based genetic algorithm and fuzzy logic. International Journal of Biomedical Engineering and Technology, 28(4), 322-348.

[55] Chen, C. H. (2018). An arrival time prediction method for bus system. IEEE Internet of Things Journal, 5(5), 4231-4232.

[56] Shankar, K., Perumal, E., & Vidhyavathi, R. M. (2020). Deep neural network with moth search optimization algorithm based detection and classification of diabetic retinopathy images. SN Applied Sciences, 2(4), 1-10.

[57] Mohanty, S. N., Ramya, K. C., Rani, S. S., Gupta, D., Shankar, K., Lakshmanaprabu, S. K., & Khanna, A. (2020). An efficient Lightweight integrated Blockchain (ELIB) model for IoT security and privacy. Future Generation Computer Systems, 102, 1027-1037.

[58] Elhoseny, M., & Shankar, K. (2020). Energy efficient optimal routing for communication in VANETs via clustering model. In Emerging Technologies for Connected Internet of Vehicles and Intelligent Transportation System Networks (pp. 1-14). Springer, Cham.

[59] Chen, C. H. (2020). A cell probe-based method for vehicle speed estimation. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 103(1), 265-267.

[60] Khamparia, A., Singh, A., Anand, D., Gupta, D., Khanna, A., Kumar, N. A., & Tan, J. (2018). A novel deep learning-based multi-model ensemble method for the prediction of neuromuscular disorders. Neural computing and applications, 1-13.

[61] Shankar, K., Zhang, Y., Liu, Y., Wu, L., & Chen, C. H. (2020). Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. IEEE Access, 8, 118164-118173.