



Detection Of Straggler Nodes Using Unsupervised Job Scheduling Method

Atul V. Dusane^{1*}, Dr. K. P. Adhiya²

Abstract:

In order to match a real work to a real resource, load balancing, distributed resource management, and matching techniques are used. By converting tasks back into distributed processing tasks, distributed data processing systems speed up work. However, rather than really charging a data cluster, relatively sluggish or straggler functions might improve system performance, delaying task completion and consuming excessive amounts of resources. Traditional straggler avoidance techniques wait to find stragglers before restarting them, which slows down straggler detection and costs money. In order to avert those situations, we are creating a system that predicts when there will be stragglers. To capture node and configuration heterogeneity, the system generated various frameworks for each network and process, which required the laborious compiling of crucial training data. In this research, we develop algorithms for multi-task learning that share information among many versions. To find the straggler, the system first gathers log data from accessible data nodes and uses an unsupervised machine learning approach. Additionally, the learning system attempts minimising such nodes rather than eliminating them after a straggler has been found. The system's ability to manage large amounts of data with few data nodes or even effectively operate in a dispersed environment with nodes that are already hot is the main benefit of this study. Unlike simplistic inter-learning frameworks, our compositions use the common context from our data, boosting the effectiveness of broad statements on limited data.

Keywords: Straggler detection, Load balancing, distributed systems, data node, maser nodes, Unsupervised learning.

DOI Number: 10.14704/Nq.2022.20.17.Nq88079

Neuroquantology 2022; 20(17):635-642

INTRODUCTION

Artificial intelligence's machine learning field gives systems the capacity to continuously learn from the past and improve on their skills without being explicitly programmed. It is necessary to purchase a machine with several TB of RAM in order to store and analyse massive amounts of data, which raises the cost of creating the programme. Another restriction is that in the event of a huge data system, it is not possible to utilise the machine learning algorithm's current methodologies or tools for analysis. Big data, or exceptionally huge data, cannot be handled by the current techniques or instruments. The first data loading might take a while because there is so much data. It's important to assign the correct resources to the relevant jobs while processing big amounts of data.

Using soft computing learning methodologies, the machine's storage capacity can manage larger amounts of data, but frequent information contact with the disc is a major worry since it degrades the application's effectiveness and speed. Programs also get more complex for the big data systems in another manner. Therefore, it becomes necessary to utilise machine learning methods, which may be used even to massive data that may be scattered in nature. The distributed systems or the data require processing so that the data dispersed across many devices or places may be combined and processed before the final processing results are supplied to the user, which can only be accomplished by the precise distribution of machine learning algorithms.

***Corresponding author:-** Atul V. Dusane

Address: - ^{1*}Research Scholar, Department of Computer Engineering, SSBT's COET, Bambhori - Jalgaon(M.S.)

²Professor, Department of Computer Engineering, SSBT's COET, Bambhori - Jalgaon (M.S.)

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.



LITERATURE SURVEY

The POC ache competitiveness caching architecture in [1] implements activity straggler tolerance. To reduce the strain associated with marginalization coding, this system divides blocks into smaller sub problems, with processes being driven by encoding procedures at the sub block level. In order to decide which similarity blocks should be cached, it additionally optimizes a straggler-aware buffer structure that takes into account both access permissions popularity and redundant estimate. The framework builds a model on Hadoop platform 3.1 HDFS in order to maintain the effectiveness and compatibility of conventional HDFS operations. In the absence of stragglers, these thorough observations on the regional and AWS EC2 clusters show that the system can cut the reading duration by up to 87.9% compared to normal HDFS. These thorough observations from both the regional and AWS EC2 clusters show that, in the absence of stragglers, POC ache may cut the reading time by up to 87.9% compared to conventional HDFS.

Applying numerical out in is the main issue with automated, responsive straggler identification on the data analytics platform [2]. The method investigated that supervised learning is legally used to adapt the best conditions to find stragglers without human interference. This investigation was motivated by recent successes over the implementation of Convolutional Neural network techniques to solve complex, efficient online challenges. According to the architecture, Hawkeye and management should develop an anomaly detection system that can launch speculative jobs on restricted devices while also identifying stragglers using reinforcement learning. Along with the findings of the analysis, Hawkeye keeps reducing the time needed to perform the task compared to other technologies of its kind. One illustration is the roughly 37% reduction in entry-level work completion date based on the advancement of 23% on the correctness of the existing resolutions to the equal-size.

A new update optimum approach for speculation strategy that automatically detects stragglers using deep learning is chosen in [3] as the overall strategy to launch speculation jobs on the network-based. Install the Croco-

dile system on a 4-node homogeneous Spark cluster test platform together with common information retrieval techniques. The research findings indicate that Cheetah reduces work project length in a variety of applications as compared to current methods, resulting in exceptional quality. As an example, it shows that compared to suggested systems on the same larger platform, average project length may be reduced by up to about 16%, and accuracy can be increased by 25%.

Compared to [4], data collection expenses would not be disregarded because machine learning often entails a heavy demand on cloud servers. In order to lessen the effects, the system employs a special software for creating and running neural networks within WSN. Object identification employing a variety of tiny, energy-efficient infrared high optical absorption with segments and sub to show this concept was one of the research the system carried out using hard numbers.

Consider the difficulty of learning the parameters of the model from data scattered over various edge devices, according to [5], in the absence of paper-provided data to a centralized place. This study appears to concentrate on a particular subset of linear regression-based computer vision models. The method assesses the decentralized logistic regression's equilibrium limit from a theoretical standpoint. The system suggests a control strategy that chooses the best balance between regional update and global averaging parameters in order to reduce the lower bound under a specified resource budget. The effectiveness of the provided approach is evaluated by exhibit demonstrations with real-world datasets, both in a networked presented technique and in a more thorough computer simulation.

The "smart" descriptions of this consumer expenditure will be forwarded to a surface framework in [6], where more advanced computer vision techniques are required to improve the precision of potential performance estimates. Three separate data sets were used to evaluate the recommended methodology, which revealed that with its discovery filter in place, it could predict failure



for up to 61.6% of cases. The good predictive classifier performance on the Surface platform was enhanced by at least an additional 70–73% by using more advanced machine learning algorithms. This innovative technology has demonstrated that it is possible to execute detection techniques in the embedded device using small and onboard artificial intelligence methods for rural operating hydrocarbon systems. A vast surveillance system is possible because to the decentralised estimate between the rural node vein wall and efficient surface based on machine learning, which delivers strong information without the need for expensive hardware or in sensing devices.

A innovative method to guard against general-purpose tampering of the distributed machine learning dataset is presented in [7]. The system then develops a mathematical model to identify the scheme's ideal training loops. The proposed technique is motivated by the observation that dispersed nodes can make it more difficult for adversaries to successfully compromise all data, even while DML may raise the danger of an attack. The system proposed a unique method to identify and rectify manipulated train samples for DML, which relied on incomplete training results, in order to guard against dataset manipulation. To guarantee data security and the accuracy of the learning model in a risky environment, the suggested strategy makes use of a cross-learning mechanism and a manipulated data detection algorithm based on variable training loops. In addition, a mathematical model is developed to identify the best training loops. Simulation tests demonstrate that the mathematical model accurately predicts simulation outcomes, with the highest performance being attained by the best training loops.

Smart surroundings are beginning to get attention in our technological era [8]. This is mostly due to the exponential growth of salaries. Wearable sensors' surroundings, jitter, and real-time nature will necessitate the use of specialized frameworks that take use of ML power optimizations in environments where dispersed inference and device communication are required. This research proposed

a centralized, interactive asset and maintain state-of-the-art neural networks, especially RNNs. The system offered performance throughout a variety of detected targets, multiple sensor nodes, data sets, and training set duration. Through hitherto undiscovered communications, this framework may be used to complete analytical tasks in a decentralized setting, according to this experimental evaluation. These findings lay a solid foundation for future studies on practise management and the effectiveness of this approach mathematically.

This behaviour has shown itself in several attacks, including DDoS. Attacker software created with the assistance of malicious individuals on IoT machines is used to carry out the assault. To battle such a scenario, it is crucial to comprehend the fundamentals of these assaults and immediately identify dangerous software. The management of homogeneous data of various sizes has also demonstrated how effective computer vision, or more particularly a subset of this technology, is. This paper suggests a method to identify problematic behaviours in IoT environments using an Artificial Neural Network. The output of this system is double that. It first distinguishes between normal and abnormal patterns of channel IoT devices, and then it develops a method with an accuracy of 77.51% for successfully preventing anomalous nodes.

According to [10] Structure for Machine Learning, a compositional net antagonistic network is used to develop an appropriate system to categorise rogue RF transceivers. The system uses a discriminating model that, when trained with transformer data in a GAN environment, prevents discrimination with 99.9% accuracy between trustworthy and fraudulent transceivers. After removing the antagonistic transmissions, the reliable transponders are then identified using a convolutional computer programme (CNN), a fully linked deep neural network (DNN), and a convolutional neural network (RNN) to demonstrate the development of a robust later part transmission recognition method with RFAL. The technique also demonstrates that for all three artificial neural networks, a higher



"trusted transmitting" identification rate is achieved whenever data from three different transmission kinds are used, as opposed to only using one type of transmission.

In order to lower the cost of deployment, a novel low-cost strategy to select a target for volatile situations is presented in [11]. By regulating the dispersed moving measurements, this technique creates an angular sense barrier that only the target may cross. The boundary deployment programme based on this profile must modify the node's depth in order to reduce migration expenses. The ideal technique considers the network size, the node's available energy, and the impact of node mobility while updating a profile. The 2-Voronoi diagrams lemma everything from nodes, and the evenly scattered triangles sink at different temperatures. When shifting nodes, that device uses experimental analysis to evaluate coverage and resource efficiency. The analysis of the observed impacts demonstrates that the method is effective. Beam irradiation can be increased and used to boost channel throughput while the cluster head is being implemented.

The focus is on the spectrum service provider network [12]. The unit's design in conjunction with a networked solution that includes a node for spectrum identification. The TV White Space (TVWS) is chosen as the transmission coverage region for the broadband programme for the scattered bandwidth detection system. Examining detection and identification methods based on the type of the radio transmitter improves the effectiveness of infrared region identification from the viewpoint of the dc generator. To improve signal prediction performance in a reduced frequency ratio framework, a phase extreme appendix approach based on radio signals I/Q data was suggested. It is proposed to distinguish between disturbance and signal using the interquartile scale. The findings of the experiment demonstrate that the electromagnetic radiation knowledge platform based on signal attributes may enhance system adaptability and detection rate.

In [13], each aircraft in a big series is regarded as a crucial node. For aviation education and groups, a new, more efficient CD approach is

being considered. Each node delivers three molecular methods from a predetermined approach pool and generates following projected tracks using an instability simulation based on a six-degree motion. The fundamental combination of tactics needed to coordinate opposing aircraft is based on the idea of maximum robustness. This work suggests a special authentication method for the arrangement of false material in order to address incomplete public consciousness. The collaboration approach to resolving upcoming data dips is dulled by this research. Two opposing situations are presented to assess the suggested fix. The findings showed a considerable potential to avoid mishaps in ongoing fights in a very complex 3-D environment with several aircraft.

A survey of the literature on the various clone node identification techniques was done in [14]. Additionally, the method supported theoretical and practical studies to categorise clone nodes in stable WSNs in light of the shortcomings and difficulties of the current decentralised and transit routes. Considerable attention is paid to wireless communications because of how widely they are used in a variety of fields, including hydrological modelling, meteorology, connected cars, location tracking, and target tracking.

DML is separated into semi-DML and ordinary DML in [15]. The dispersed computers are then sent educational activities from the central server and learning outcomes are then consolidated. The central server also acknowledges and compensates users based on their understanding of databases and how they work in straightforward DML. The programme first proposed a simpler analytical DML Data Poisoning Detection System, which employs a boundary technique to detect the toxic data. According to the procedure, a computer algorithm is used to create training circuits that the power border firm then uses to develop circuits with the optimal number of repeats. Instead, the tool improves the DML's awareness of data poison, allowing for stronger learning protection through the primary resource. An operating system's best utilisation requires the development of an effective resource allocation strategy. Accor-



ding to computation findings, the enhanced plan will significantly improve the entire prototype's outcomes by up to 20% for the help neural network and by 60% for the straightforward DML logistic regression. Additionally, the enhanced information poison detection system with optimum resource allocation would save 20% more time and money in the semi-DML situation.

An application that can deliver on-demand service requests is presented in a fog node at [16] intelligence agency. As virtual machines are resource restricted, our objective is to incorporate mild VMware vsphere, such as Containers, to improve the two electrodes framework. On the fog networks that were simulated by a Raspberry Pi-4 device, the device was able to run a number of software security instances for intrusion detection. Several system flows are being examined in this setting to confirm its efficacy. The improvisation approach for the safety implementer was established as a collection of guidelines for practical, delicate harmonies in IoT with visibility in this cloud methodology. The results of this study ensure the creation of a framework for pollution management that promotes the development of safer IoT devices.

Focuses on locating and resolving defects, state estimate, and condition monitoring, according to [17]. There is a substantial amount of information utilised for condition monitoring. This programme will be utilised for classification jobs, and it is feasible to save that routing information between communication nodes. Data mining methods must also be provided owing to the features of the supplied data. The maintenance activities research has the benefit of minimising expenses, which is the major driver behind any newly founded organisation. This enables the results to be applied to other businesses in order to solve issues in the future well in advance.

The M2M vehicle-based Machine Learning Network (ML) based trust system [18] focuses on the modelling technique for suspicious behaviour identification (MLBT). Additionally to the Intense Recharge Entropy Dependent Data Augmentation (EBFE), a component of the

battery (XGBoost) technology, is programmed into the system using Binary Based Optimization methods. The performance of this approach is evaluated using three performances with various confusion matrices, following state-of-the-art distinct classifiers like XGBoost and Randomized Forestry. With a 10% improvement in accuracy, TPR, and FPR compared to the invader density state-of-the-art designs, the simulated results show the recommended model's effectiveness.

Recognizing fake relay nodes in a mesh network relies on unmonitored information based on pattern results gathered from the Supplier connection at the self-definition [19]. The system must accurately represent the actual conditions under which aggressive communications system intrusions, contaminating agents, or slurring attacks—which focus primarily on receiver antenna—could occur. This kind of threat detection problem is addressed by applying unprotected data science and one-class identification methods. Kernel-based radial components, in contrast to architectures, offer the best accuracy for identifying malicious network assaults with these kinds and for finding precise relays by using the distinctive symbol configuration aspects of the transmitted signal. Results show that for muttering relay attacks, SVM-RBF and k-NN machine learning can reach detection accuracy of almost 99%. The findings also support OCC computers' recognition of various relay attack methods, each of which is viewed with specific feature preferences in this investigation.

In [20], a computer technology to permit certain entities within a sparse networks system was proposed. Here, the method was utilised to determine the sensor traffic based on past data gathered by all of these local controllers.

PROPOSED SYSTEM DESIGN

The general strategy for finding the lone survivors in large data structures. By merging the characteristics of the structure and the procedure, the system is able to pinpoint the reasons for a larger spectrum for the stragglers. To reduce false positive findings from the root cause analysis, they integrate



statistical principles for the various features. We make use of edge detection to prevent excessive resource utilisation on the primary job, which is the major reason stragglers occur, and we estimate inferior bounds using empirical ones. The traits are regarded as the primary factor.

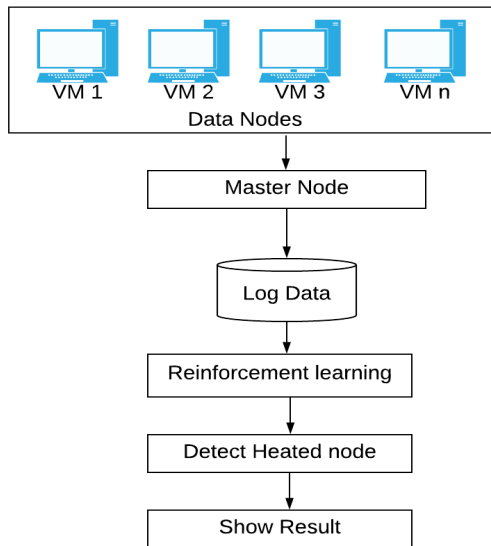


Figure 1: Process of detection Straggler Nodes,

The outcome of the whole process is the detection of straggler nodes (Figure 1), which the system then automatically stops. The set of actions below should be adhered to during execution. Before allocating any node, we collect the data locality characteristics, CPU load, and memory load from all virtual machines before submitting the different workloads to that node. This data is taken into consideration as load data to locate the straggler during execution.

ALGORITHM DESIGN

Algorithm: Policy based Q- Learning Algorithm

Input: $Input[1.....n]$ total input parameters that is generated by different input objects, set of Threshold values $T_Min[1.....n]$ and $T_Max[1.....n]$ for all inputs.

Output: Activate the message as node is straggler or not.

Step 1: Extract all parameters instance from DB (Row from Database)

Step 2: $Params_Parts [] \leftarrow Splitting(Row)$ using below formula

Step 3: $Par_Parts [] = \sum_{k=0}^n Col[k]$

Step 4: Validate (C_val with given threshold set of $T_Min[1.....n]$ to $T_Max[1.....n]$)

If (condition is true)

Perform activate on respective output appliances.

Else Continue;

Step 5: $Timestamp(tm) \leftarrow Extract\ current\ VM\ time.$

Step 6: if ($Timestamp_tm > Bount_Time$)

Give penalty to all measures $TrueP++$ and reward $FalseN++$

Otherwise continue. $Total++$

Step 7: compute penalty $weight = (TrueP * 100 / Total)$

Step 8: if ($weight \geq Thval$)

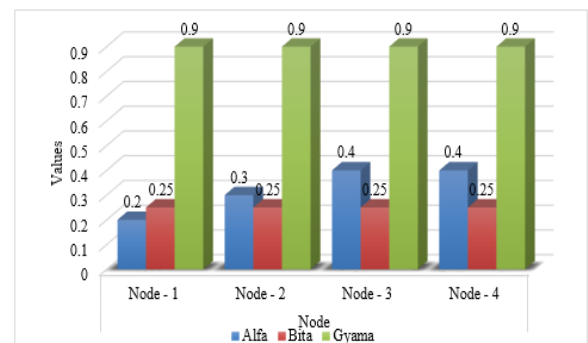
Predict straggler event

RESULTS AND DISCUSSION

Four data nodes and one master node are part of the system's open-source local distribution architecture. Using random task creation and a dynamic job ordering approach, the workload was produced dynamically for each individual data node. All of the data nodes have uniform configuration settings, and they all read different node characteristics such as CPU load, memory load, and data locality information before being used to train supervised machine learning classifiers. The system has been verified using an unsupervised learning approach, as shown in Table 1 and Graph 1 below.

Table 1: Configuration parameter setting to each data node

Unsupervised learning Method	Node			
	Node - 1	Node - 2	Node - 3	Node - 4
Alfa	0.2	0.3	0.4	0.4
Bitu	0.25	0.25	0.25	0.25
Gyama	0.9	0.9	0.9	0.9

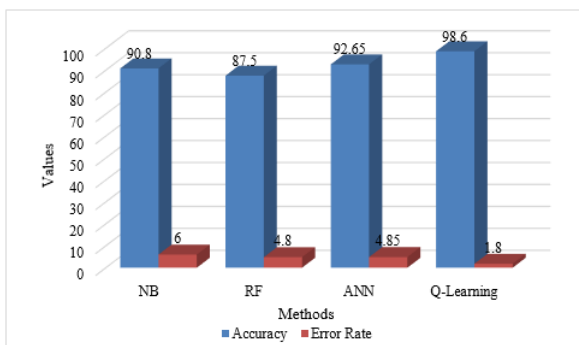


Graph 1: Configuration parameter setting to each data node.



Table 2: Straggler identification utilising suggested unsupervised learning classifier and supervised learning classifier with high classification accuracy

Unsupervised learning and supervised learning methods	Accuracy	Error Rate
NB	90.8	6
RF	87.5	4.8
ANN	92.65	4.85
Q-Learning	98.6	1.80



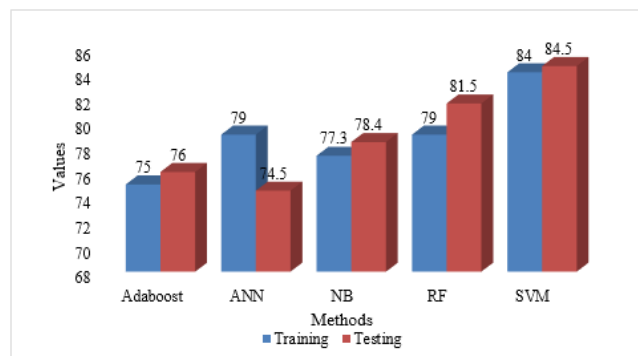
Graph 2: Straggler identification utilising suggested unsupervised learning classifier and supervised learning classifier with high classification accuracy

The accuracy of straggler node detection utilising a variety of supervised machine learning methods, including Naive Bayes (NB), Random Forest (RF), Artificial Neural Network (ANN), and the proposed Q-Learning-based unsupervised machine learning classifier, is shown in Graph 3 above. Compared to conventional machine learning techniques, the suggested approach offers a higher detection accuracy. The accuracy of straggler node detection utilising a variety of supervised machine learning methods, including Naive Bayes (NB), Random Forest (RF), Artificial Neural Network (ANN), and the proposed Q-Learning-based unsupervised machine learning classifier, is shown in Graph 3 above. Compared to conventional machine learning techniques, the suggested approach offers improved detection accuracy.

Another machine learning experiment that uses categorization demonstrates the system's performance. The distribution of the whole dataset's train and test data sets is shown in Graph 3 below; each area displays the proportion of occurrences of the dataset. Different cross-fold data validation methods, such as 5, 10, and 15, have been employed with the traditional 70–30% data splitting methodology.

Table 3: Data splitting with various cross-validations for training and testing.

Methods	Training	Testing
Adaboost	75	76
ANN	79	74.5
NB	77.3	78.4
RF	79	81.5
SVM	84	84.5



Graph 3: Data splitting with various cross-validations for training and testing.

The detection of straggler nodes on the distributed log dataset has been done using five supervised classification techniques, as shown in graph 3 above. In comparison to other classifiers, the ANN has lower accuracy than the SVM. We used a different dataset as test examples produced by several log data sources to evaluate the effectiveness of such classification systems. Performance evaluations have been conducted using a variety of current systems that have created approaches using various machine learning algorithms. A separate experiment was conducted with several data samples, and the classification accuracy was afterwards tested. The results of the experiment in Table 1 clearly shown that the suggested course has a high quality, using just two training classes, as opposed to the usual grouping of the proposed algorithm, which uses different training classes. The system has been compared to the proposed system and assessed using random cross-validation throughout the whole file system.

CONCLUSION AND FUTURE WORK

Big data is made up of a variety of knowledge that may be found in nature. Smaller parts of this data that are divided up for compilation are referred to as jobs. The suggested designs are useful for acquiring relevant data log from virtual computers. They divide up the duties or smaller jobs into smaller jobs or portions. This



phenomenon, known as an outlier in some circumstances, causes several studies to become slower than other activities, which slows down the foundation of the entire programme. Outliers have a negative impact on the effectiveness of huge data systems.

High CPU consumption, high disc utilisation, network package failure, hardware problems, data skew, etc. are always factors in the occurrence of outliers. With this method, the traits of outliers and regular data are compared simultaneously. This feature is considered the root cause of outliers if the relevance of a feature of exceptions significantly differs from that of a typical activity. Future work on this system will be fascinating since it will include working with huge distributed data nodes or virtual computers in a distributed environment.

REFERENCE

- Zhang, Mi, et al. "Parity-Only Caching for Robust Straggler Tolerance." 2019 35th Symposium on Mass Storage Systems and Technologies (MSST).IEEE, 2019.
- Du, Haizhou, and Shaohua Zhang. "Hawkeye: Adaptive Straggler Identification on Heterogeneous Spark Cluster with Reinforcement Learning." IEEE Access 8 (2020): 57822-57832.
- Du, Haizhou, et al. "Cheetah: A dynamic performance optimization approach on heterogeneous big data analytics cluster." 2019 5th International Conference on Big Data Computing and Communications (BIGCOM).IEEE, 2019.
- Yamaguchi, Hirozumi. "Distributed Machine Learning Over Wireless Sensor Networks." 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops).IEEE, 2019.
- Wang, Shiqiang, et al. "Adaptive federated learning in resource constrained edge computing systems." IEEE Journal on Selected Areas in Communications 37.6 (2019): 1205-1221.
- Greff, Heloise, et al. "Distributed inference condition monitoring system for rural infrastructure in the developing world." IEEE Sensors Journal 19.5 (2018): 1820-1828.
- Chen, Yijin, et al. "Data Manipulation Avoidance Schemes for Distributed Machine Learning." ICC 2019-2019 IEEE International Conference on Communications (ICC).IEEE, 2019.
- Abudu, Prince, and Andrew Markham. "Distributed Communicating Neural Network Architecture for Smart Environments." 2019 IEEE International Conference on Smart Computing (SMARTCOMP).IEEE, 2019.
- Khatun, MirzaAkhi, NiazChowdhury, and Mohammed NasirUddin. "Malicious Nodes Detection based on Artificial Neural Network in IoT Environments." 2019 22nd International Conference on Computer and Information Technology (ICCIT).IEEE, 2019.
- Roy, Debashri, et al. "RFAL: Adversarial Learning for RF Transmitter Identification and Classification." IEEE Transactions on Cognitive Communications and Networking (2019).
- Sun, Jiayi, and Gaotao Shi. "Cost-Efficient Node Deployment for Intrusion Detection in Underwater Sensor Networks." 2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS).IEEE, 2019.
- Chen, Zhenjia, and Yonghui Zhang. "Providing Spectrum Information Service Using TV White Space via Distributed Detection System." IEEE Transactions on Vehicular Technology 68.8 (2019): 7655-7667.
- Wan, Yu, Jun Tang, and Songyang Lao. "Distributed conflict-detection and resolution algorithms for multiple uavs based on key-node selection and strategy coordination." IEEE Access 7 (2019): 42846-42858.
- Numan, Muhammad, et al. "A systematic review on clone node detection in static wireless sensor networks." IEEE Access 8 (2020): 65450-65461.
- Chen, Yijin, et al. "Data Poison Detection Schemes for Distributed Machine Learning." IEEE Access 8 (2019): 7442-7454.
- Imrith, Vashish N., et al. "Dynamic Orchestration of Security Services at Fog Nodes for 5G IoT." ICC 2020-2020 IEEE International Conference on Communications (ICC).IEEE, 2020.
- Dsouza, Joanita, and SenthilVelan. "Preventive Maintenance for Fault Detection in Transfer Nodes using Machine Learning." 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE).IEEE, 2019.
- Eziama, Elvin, et al. "Detection of Adversary Nodes in Machine-To-Machine Communication Using Machine Learning Based Trust Model." 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT).IEEE, 2019.
- Yengi, Yeliz, et al. "Malicious relay node detection with unsupervised learning in amplify-forward cooperative networks." 2019 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT).IEEE, 2019.
- Jaint, Bhavnesh, et al. "Malicious Node Detection in Wireless Sensor Networks Using Support Vector Machine." 2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE).IEEE, 2019.

