

Identifying Clone Nodes in Wireless Sensor Networks with Minimal Communication and Storage Capacity Requirements

Charanjeet Singh
Assistant Professor

Electronics and Communication
Department
Deenbandhu Chhotu Ram University of
Science and Technology
Murthal
charanjeet.research@gmail.com

Akshay Rajput
Assistant Professor

Department of Computer Science &
Engineering
Graphic Era Deemed to be University
Dehradun, Uttarakhand, India
akshay.rajput@geu.ac.in

Ajay Rana
Amity University

Greater Noida
Uttar Pradesh
India
ajay_rana@amity.edu

Biruk Yirga

CSE Department
Adama Science and Technology
University (ASTU)
Adama
Ethiopia
birukyirga73@gmail.com

Dr. A Y Prabhakar
Professor

Department of ETC
Bharati Vidyapeeth Deemed to be
University College of Engineering
Pune, Maharashtra, India.
ayprabhakar@bvucoep.edu.in

Ashish Parmar

Lloyd Institute of Engineering and
Technology
Greater Noida
Uttar Pradesh, India.
apsvgi@gmail.com

Abstract- The "sensor nodes" in a wireless sensor network are the individual devices that collect data and are powered independently. These nodes act as sensors and processors. The credentials of a compromised node can be used to create multiple identical "clone" nodes elsewhere in the network. The community could be harmed in the short or long term by the proliferation of these clone nodes. In this paper, the authors discuss a hashing-based technique that avoids the need for this information. Every cluster's leader implements the proposed method for locating clone nodes, considering the hierarchy present in the data. The simulation results demonstrate its superiority over RDB-R and TDS.

Keywords – Clone node or duplicated node, Wireless Sensor Network Computational complexity, Probability of detection, Storage overhead

I. INTRODUCTION

Wireless sensor networks are characterised by nodes that are completely self-contained, generate their own power, and work together to achieve a shared objective (WSNs). These nodes, when distributed throughout an area, may gather, analyse, and send data in order to carry out a variety of functions, including environmental monitoring and the detection of fires. WSNs are utilised rather regularly for a variety of purposes, including but not limited to the following: monitoring the health of machines; monitoring industrial operations; researching marine ecosystems; exploring hazardous regions; monitoring caves; and monitoring mining activities. WSN is utilised for a broad number of monitoring objectives, some of which include monitoring the health of machines, the health of the military, the health of the industrial sector,

and the health of individual consumers. [1] [2]. WSN is susceptible to a wide number of different types of assaults. Attacks may be classified into two distinct groups: those that call for two tiers of defence, and those that don't. [3] [4].

In the case of an assault, the person responsible for it may be situated anywhere, either inside or outside of the network. If hackers already have access to the network, they may try to get access via analysing the network's structure. On the other hand, attackers coming from outside the network are not truly connected to it in any way. An illustration of a layer-agnostic attack is a clone assault, which is sometimes referred to as a node replication attack. An attacker has the ability to physically grab a sensor node, retrieve data such as the node's id, and then return to the network to duplicate more sensors with the credentials that they have obtained. The cloned version retains the original sensor's distinct identification as well as its fundamental properties and characteristics. This information will be utilised by a cloned node in order to successfully establish a connection to the network. These nodes have the potential to cause serious problems for the network as a whole. These nodes seed the cloned node's network traffic with manufactured information while also monitoring network traffic through the copied node.

Damage might be done to the network as well as the information that is being delivered if there is a clone node present. We have conducted the necessary investigation and come up with a few potential solutions to the problem. There are several different approaches that may be taken to

target clones in wireless sensor networks that are static. On the other hand, the bulk of these methods are dependent on either information from GPS or sensors located nearby. There are now two ways available for identifying clone nodes, and they are referred to be centralised and decentralised. Due to the many drawbacks associated with methods with a single center, people rarely use centralised methods. The base station is responsible for data collection and analysis across the network in order to identify any potential clone nodes. Having stated that, the base station is the component that is the utmost important. There is no way that detection can work properly in the absence of a working base station. The adoption of decentralised strategies has resulted in a shift away from concentrating attention on a central location. The discovery of clones requires participation from all of the nodes. Decentralized methods have become increasingly popular in recent years as a response to the drawbacks of centralised ones.

Look down below to see how the rest of this project is put together! In this second and final part of the series, we will talk about the extensive literature survey that was done in this area. The research provides in-depth descriptions of both centralised and decentralised methods of operation. [Citation needed] [Citation needed] The nature of the problem and the strategy that was used to implement the network are broken down in detail in the third section of this analysis of a network. In the fourth part of this breakdown, the approach that has been suggested has its underlying algorithm outlined. We calculate the complexity, and then we show how likely it is that the hash will collide with itself. Section 5 contains a presentation of the results of the simulation. In this section, we will talk about the simulation's parameters and make comparisons between the two different simulation runs. The project is summed up in Section 6, which also takes a look into the future of scientific investigation.

II. LITERATURE SURVEY

According to the theory of Parno et al. [5,] every node is able to interact with its surrounding nodes. This enables detection to be performed centrally. The sink node is able to determine the identities of the clone nodes since it is aware of the locations of those nodes. The random key pre distribution and the bloom filter are the two fundamental components of the new method that Brook et al. [6] have presented. The fact that it has such a high percentage of false positives and false negatives is its greatest shortcoming.

Parno et al. [5] present a method that demonstrates how trustworthy multicasting may be accomplished through the utilisation of "witness nodes." An assault against this witness node is something that can be attempted. Parno et al. [5] also highlighted the several ways that RM and LSM can be located. Within RM, assertions of location are transmitted to "witness" nodes that have been chosen at random. LSM, in contrast to RM, makes it simple to

pinpoint a particular forwarding node since every node keeps a record of its own claimed position. This makes it possible for LSM to replace RM. Zhu et al. [7] demonstrated the existence of two distinct forms of LSM: SDC and P-MPC. The manner in which you go about doing this will also be impacted by where you are. M. Conti and his fellow employees [8] developed RED, which can only be used within specific time slots. There are usually two steps involved in the execution of a protocol. The first thing that occurs is that a random number with the designation rand is transmitted from the base station to each of the nodes in the network. We will search for anything during the second step, which we will refer to as "detecting." During the phase of detection, every node in the network broadcasts its claim in order to connect with the other nodes that are close (ID and position). Several witness nodes will be contacted at various points over the duration of the execution. This method needs little memory but effectively zeroes in on the problem at hand. The location is another factor that contributes to the overall effectiveness of this technique. A. K. Mishra [9] is the one who came up with the idea for the zone-based node replica detection technique. This tactic is effective because it divides the network into a number of zones, each of which is led by a leader who is responsible for identifying instances of duplicate nodes. Within the framework of this protocol, the finding of information occurs in two steps. The process of detection is broken up into two stages: first, it takes place within a certain zone, and then it moves on to take place between zones. Unsecured wireless networks are susceptible to attack if an unauthorised user successfully impersonates the zone leader. Kwatae Cho et al. detail one approach to replica discovery that does not rely on geolocation data but rather on the identifiers of neighbouring nodes in their paper [10]. This approach to replica discovery is described in more detail. Through the use of a bloom filter, the IDs of neighbouring nodes are equally decreased. Using the IDs of their neighbours, our method uncovered any clone nodes in the network. A bloom filter is applied so that node IDs that are adjacent to one another look the same size. Bloom filter output, also known as BFO, is sent to particular nodes so that evidence can be gathered. The Randomized Distributed Bloom filter using Replica (RDB-R) scheme performs exceptionally well. This is due to the fact that it is not dependent on any particular deployment strategy and does not call for a connection to a global positioning system (GPS). This leads to significantly lower costs for the sensor nodes themselves as well as increased communication capacity. Despite this, there are problems with the level of detection as well as the amount of memory that is needed. Zhimingzhang et al. [11] propose a TDS with orthogonal code, which is comparable to [12 - 17].

Calculations of energy are necessary for every authentication-based method, which are described in [18][19][21][22][23][24]. [18][19][21][22][23][24] In [21], it is discussed how to detect a node replication attack in mobile wireless sensor networks by tracking the


```

Step 1: Set B[] ← 0, Sumb ← 0

Step 2: For each node k in the cluster
        Step 3: j ← k mod N

Step 4: Increment B[j]

Step 5: For i ← 0 to N-1

Sumb ← Sumb + B[i]

Step 6: If Sumb ≠ Suma
        For i ← 0 to N-1

                If A[i] ≠ B[i]

                        If B[i] - A[i] < Ti send an alert
                        message to all clusters
                Else

                        Discard all nodes that mapped to
                        A[i] from the network.
    
```

Figure 3 illustrates the application of the LCLS algorithm. In Figure 3, we can see Cluster X, which consists of seven nodes, constructing its list A and calculating Sum_a by collecting messages (a). Following the recording of each value in list B, that value is then compared to the entry in list A that most accurately characterises it. When the difference between B and A reaches a certain level, the neighbouring clusters are notified about the situation. Only the nodes that are still linked to the same physical location are left when others in a cluster are removed. Figure 3 demonstrates how the node 1345 is replicated across all of the nodes that make up cluster x. (b). A comparison of lists A and B is used to generate warning signals that are sent to clusters that are located in close proximity to one another.

The time that is necessary to complete the LCLS algorithm that has been presented stays the same regardless of how long the list is. The maximum size of the list is directly proportional to the total number of nodes in the cluster. Therefore, the amount of work that is required is $N \log N$, where N is the total number of nodes that are part of the cluster. This operation is carried out C times whenever there are C network clusters.

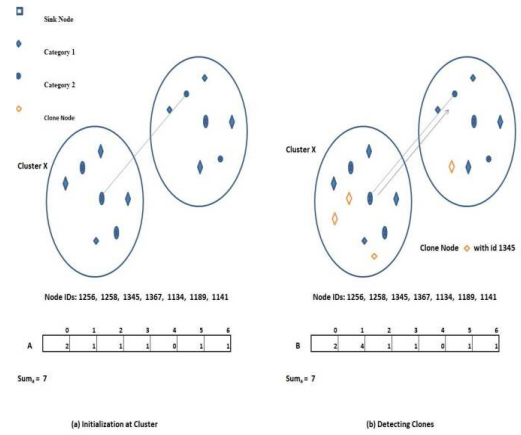


Fig. 3 LCLS

The overall computational complexity may be calculated using the formula $T = O(C * N)$. When using the proposed technique, the length of time it takes to transmit a message is determined by the number of degrees possessed by each node. Every cluster head that discovers a clone node quickly notifies the other heads in its neighbourhood. The eliminated nodes make it more difficult for the cluster head to see the whole network. You are able to examine the relative prices of communication and storage in Table 1.

TABLE 1 COMPARISON OF COMMUNICATION AND STORAGE COST

Algorithm)	Cost of communication	Cost of storage
RM	$O(n)$	$O(n)$
LSM Random Key Distribution	$O(\sqrt{n})$ $O(gpdn\sqrt{n})$ $O(\log n)$	$O(\sqrt{n})$ $O(gpd)$ $O(k)$
RED TDS	$O(d)$	$O(d+M)$
Proposed methodology of -LCLS	$O(d)$	$O(N+d)$

V. RESULTS AND DISCUSSIONS

On the basis of the information shown in Table 2, we can see that the region that was analysed has a dimension of 500 by 500 pixels, with 200 nodes falling into Category 2 and 100 nodes falling into Category 3. The simulated network is an unmoving wireless sensor network. If the range is set to the value specified, each node has a detection range of $0.523m^2$, which results in a total detecting and transmission range of 0.5 metres throughout the network.

A. Probability of Clone Detection

Clone nodes in a wireless sensor network (WSN) might give rise to a variety of problems and even intrusion attempts. The probability may be determined by taking the average of the outcomes of several simulations. We have come up with a technique that can identify duplicate nodes far more quickly than TDS can. A comparison of the likelihood of detection may be seen in Figure 4.

$$P = \frac{\text{Number of Clone nodes detected}}{\text{Total number of Clone nodes deployed}} \quad (1)$$

17-32	32 x 32
33 - 64	64 x 64
65 - 128	128 x 128

TABLE 2 PARAMETERS USED IN SIMULATION

Field area	500*500
Number of Category 1 Sensor nodes	100
Number of Category 2 Sensor nodes	200
Sensor node's position	Fixed
Transmission range:	
Category 1	1m
Category 2	5m
Antenna	Omni directional

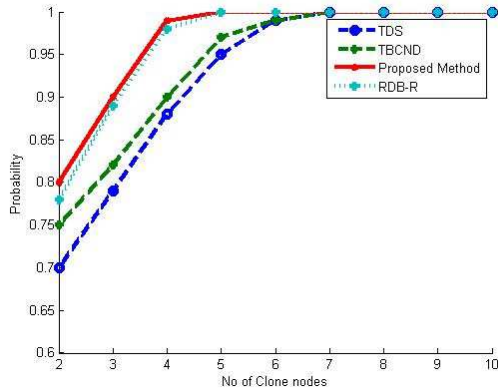


Fig. 4 . Probability of Clone Detection

Depending on the value of the new threshold, the number of clone nodes may either rise or decrease. Before continuing with replica discovery, the leader of the cluster has to make sure that any nodes that are currently asleep are taken into consideration. The position of the sleeping node in the list is decreased when the cluster master does the calculation (id / N). When a new worker node is added to the network, the cluster master will increase the total by one after performing an update.

B. Overhead Storage

In this section, we conduct an analysis of TDS in relation to the storage load. The storage load is anticipated due to the additional space required to store not just the data and code that are produced by the method itself, but also the data and code that are generated by the method itself. When developing WSN procedures, some level of thought must be given to the amount of memory that is required for each step. Memory requirements for the TDS orthogonal matrix are presented in Table 3, which reveals their considerable size.

TABLE 3 AN ORTHOGONAL MATRIX IS NECESSARY.

No. of Nodes	Orthogonal Matrix size
1-2	2 x 2
3-4	4 x 4
5-16	16 x 16

Our plan calls for the utilisation of a single list, the size of which is directly related to the number of nodes. According to Table 3, the TDS algorithm requires a 128-by-128 matrix in order to create an orthogonal code for a 100-node sensor network. Because of this, the total amount of RAM that is required is 128×128 bytes. Each node would have storage for a code that is 128 bytes long. The use of this code makes it easier for parties to communicate with one another. The LCLS, on the other hand, does not require either code or other information. In addition to this, it only needs N bytes for each cluster head, where N is the total size of the cluster[30].

When modelling LCLS, we used a broad variety of sensor densities and cluster sizes in our simulations. A typical sensor cluster had anywhere from 20 to 40 sensors, and the overall number of sensors varied anywhere from 400 to 900. Figure 5 illustrates how the activities that were taken turned out as a result. The minimum amount of random access memory (RAM) required by each cluster head will decrease as the cluster grows in size.

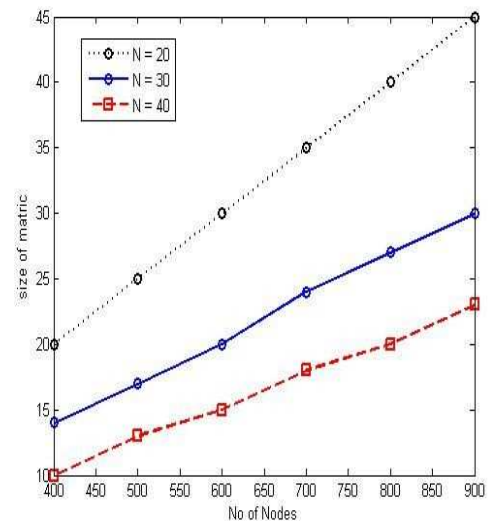


Fig. 5 overhead Storage

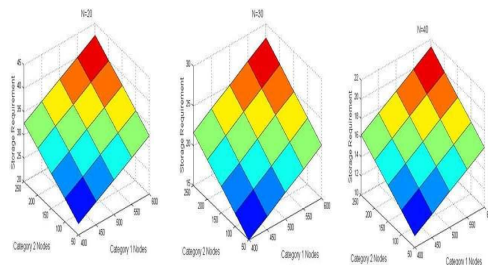


Fig. 6 The storage overhead as the number of nodes increases or decreases

Figure 6 illustrates the need for additional space for the storage of data in clusters that contain a variety of nodes. In

order to investigate this, we ran our simulations fifty times, once with $N=20$, once with $N=30$, and once with $N=40$. In contrast, the number of type 2 nodes ranges from 50 to 250, with type 1 sensors totaling between 400 and 600. The amount of storage overhead (in bytes) at the network's cluster head is proportional to the number of nodes in each cluster[31]. This is the case regardless of whether the nodes are type 1 or type 2.

C. Computational Overhead

The information that has been conveyed to the cluster leader in the form of messages is utilised in the process of creating the list. The amount of time that the proposed method requires, in comparison to RDB-R and TDS, is presented very clearly in Figure 7. We looked at a total of 500 vertices in the graph[32]. The number of clone nodes produced ranged anywhere from 20 to 45. Equation 2 from De Meulenaer and colleagues was what we used in order to figure out an object's energy level. [14]

$$E_{CP} = E * \text{Clock Cycles} \quad (2)$$

E_{CP} is computational overhead and E is energy required per clock cycle.

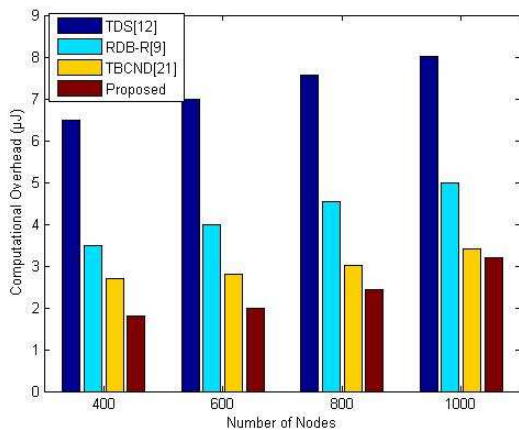


Fig. 7 overhead comparison

VI. CONCLUSION

In this study, we proposed an alternate method for locating and removing clones in a WSN that uses mesh networking technology.[33] The strategy that has been suggested focuses on finding a clone node in a manner that does not call for the collection of location data or the writing of new code for nodes. The proposed process is modelled in this study, and the results and their explanation are provided[34]. This strategy of finding clone nodes that make connections to cluster nodes from outside the cluster zone has space for improvement, though.

V. REFERENCES

- [1] I.F. Akyildiz, *et al.*, "A Survey on Sensor Networks," in *IEEE Commun. Mag.*, vol. 40, no. 8, pp.102-114, Aug, 2002.
- [2] J. Deng, *et al.*, "A Survey on Sensor Networks," in *security, privacy, and fault tolerance in wireless sensor networks*. Artech House, August 2005.
- [3] T. Bonaci, *et al.*, "Distributed clone detection in wireless sensor networks: an optimization approach," in *Proceedings of the 2nd IEEE International Workshop on Data Security and Privacy in Wireless Networks (WoWMoM '11)*, Lucca, Italy., June 2011.
- [4] Patil, D. D., Singh, A. K., Shrivastava, A., & Bairagi, D. (2023). IOT Sensor-Based Smart Agriculture Using Agro-robot. In *IoT Based Smart Applications* (pp. 345-361). Springer, Cham.
- [5] Borole, Y. D., Shrivastava, A., & Niranjanamurthy, M. (2022). Diagnosis of COVID-19 Using Low-Energy IoT-Enabled System. *IoT Based Smart Applications*, 375.
- [6] Saxena, M. C., Banu, F., Shrivastava, A., Thyagaraj, M., & Upadhyay, S. (2022). Comprehensive analysis of energy efficient secure routing protocol over sensor network. *Materials Today: Proceedings*.
- [7] Haripriya, D., Kumar, K., Shrivastava, A., Al-Khafaji, H. M. R., Moyal, V., & Singh, S. K. (2022). Energy-Efficient UART Design on FPGA Using Dynamic Voltage Scaling for Green Communication in Industrial Sector. *Wireless Communications and Mobile Computing*, 2022.
- [8] Shrivastava, A., Rizwan, A., Kumar, N. S., Saravanakumar, R., Dhanoa, I. S., Bhambri, P., & Singh, B. K. (2021). VLSI implementation of green computing control unit on Zynq FPGA for green communication. *Wireless Communications and Mobile Computing*, 2021.
- [9] W. T. Zhu *et al.*, "Detecting node replication attacks in wireless sensor networks: a survey," in: *a survey, Journal of Network and Computer Applications*, vol. 35, no. 3, pp. 1022–1034, 2012.
- [10] B. Parno, *et al.*, "Distributed detection of node replication attacks in sensor networks," in *In Security and Privacy, 2005 IEEE Symposium*. pages 49 -63, may 2005.
- [11] Brooks R, *et al.*, "On the detection of clones in sensor network using random key pre distribution," in *IEEE Trans. Syst. Man. Cybern.* vol .37 No. 6, pp. 1246-125, 2007.
- [12] Bio Zhu, *et al.*, "Localized multicast: efficient and distributed replica detection in largescale sensor networks," in *IEEE Transactions on Mobile Computing*, Vol. 9, No. 7, pp 913-926, July 2010.
- [13] M. Conti, *et al.*, "A Distributed detection of clone attacks in wireless sensor networks," in *IEEE Transactions on Dependable and Secure Computing*, vol. 8, no. 5, pp. 685-698, September/October 2011.
- [14] A. K. Mishra and A. K. Turuk, "A zone-based node replica detection scheme for wireless sensor networks," in *Springer, Wireless personal communications*, vol. 69, no. 2, pp. 601-621, 2013.
- [15] Kwantae Cho, *et al.*, "A zone-based node replica detection scheme for wireless sensor networks," *Low-Priced and Energy Efficient Detection of Replicas for Wireless Sensor Networks*, Vol. 11, NO. 5, September/October 2014.
- [16] Zhiming Zhang, *et al.*, "An efficient detection scheme of node replication attacks for wireless sensor networks," in *International Journal of Security and Networks*, VOL. 10, NO. 4, MAY 2015.
- [17] Chia-Mu Yu, *et al.*, "Compressed Sensing-Based Clone Identification in Sensor Networks," in *IEEE Transactions on Wireless Communications*, VOL. 15, NO. 4, APRIL 2016.
- [18] Vandana Mohindu, and Yashwant Singh, "Node authentication algorithm for securing static wireless sensor networks from node clone attack," in *International Journal of information and computer security*, Vol. 10, No 2/3, 2018.
- [19] De Meulenaer, G., *etal.*, "On the energy cost of communication and cryptography in wireless sensor network," in *IEEE international Conference on Wireless and Mobile Computing Networking and Communications, WIMOB '08, IEEE* October, pp. 580-585, 2008.
- [20] H. Choi, S. Zhu, and T.F. L.Porta, "SET: Detecting clones in sensor networks," in *Proc. Security Privacy Communication Network, Workshops*, pp. 341-350, 2007.

- [21] K wantae Cho, *et al.*,"Classification and experimental Analysis for Clone Detection Approaches in Wireless Sensor Networks," in *IEEE Sys. Journal*, vol 7, no. 1 Mar. 2013.
- [22] Sathish R and Kumar D R,"Dynamic Detection of Clone Attacks in Wireless Sensor Networks," *International Conference on Communication Systems and Network Technologies(CSNT)*, pp.501-505, 6-8 April 2013.
- [23] K. Farah and L. Nabila,"The MCD Protocol for Securing Wireless Sensor Networks against Node Replication Attacks," *International Conference on Advanced Networking Distributed Systems and Applications*, Bejaia, pp. 58-63,2014.
- [24] Soderlund, R., *et al.*,"Energy efficient authentication in wireless sensor networks, IEEE Conference on Emerging Technologies and Factory Automation," in *IEEE Conference on Emerging Technologies and Factory Automation, ETFA, IEEE*, September, pp. 14121416, 2007.
- [25] Wendi RabinerHeinzelman, *et al.*,"Energy- Efficient Communication Protocol for wireless micro sensor networks," in *33rd IEEE international Conference on system sciences*, 2000.
- [26] Sachin Lalar, *et al.*,"An efficient tree based clone detection scheme in wireless sensor networks," in *Journal of Information and Optimization Sciences*, vol. 40, pp. 1003-1023, 2019.
- [27] Mohamed Elshrkawey and M. ElsayedWahed,"An Enhancement Approach for Reducing the Energy Consumption in Wireless Sensor Networks," *Journal of King Saud University- Computer and Information Sciences*, vol. 30 issue 2, pp. 259-267 April 2018.
- [28] Bhupendra and VidushiSharma,"Energy efficient communication overhead algorithm in wireless sensor networks," *3rd IEEE International Advance Computing Conference*,2223 February, 2013.
- [29] Segun O Olatinwo and Trudi H Joubert,"Efficient energy resource utilization in a wireless sensor system for monitoring water quality," *Journal on wireless Communication and Networking*, January 2019.
- [30] Kunwar, V., Agarwal, N., Rana, A., Pandey, J.P. (2018). Load Balancing in Cloud—A Systematic Review. In: Aggarwal, V., Bhatnagar, V., Mishra, D. (eds) Big Data Analytics. Advances in Intelligent Systems and Computing, vol 654. Springer, Singapore. https://doi.org/10.1007/978-981-10-6620-7_56
- [31] Priyanka Chawla, Inderveer Chana, Ajay Rana, Cloud-based automatic test data generation framework, *Journal of Computer and System Sciences*, Volume 82, Issue 5, 2016, Pages 712-738, ISSN 0022-0000, <https://doi.org/10.1016/j.jcss.2015.12.001>.
- [32] Gupta, S., Rana, A., Kansal, V. (2020). Optimization in Wireless Sensor Network Using Soft Computing. In: Raju, K., Govardhan, A., Rani, B., Sridevi, R., Murty, M. (eds) Proceedings of the Third International Conference on Computational Intelligence and Informatics . Advances in Intelligent Systems and Computing, vol 1090. Springer, Singapore. https://doi.org/10.1007/978-981-15-1480-7_74
- [33] Azad, Murari Lal, Shubhranshu Vikram Singh, and Aizad Khursheed. "Improving Voltage profile of a grid, connected to wind farm using static var compensator." *International Journal of Advances in Engineering & Technology* 7.5 (2014): 1497.
- [34] Singh, S. Vikram, Aizad Khursheed, and Zahoor Alam. "Wired Communication Technologies and Networks for Smart Grid—A Review." *Cyber Security in Intelligent Computing and Communications* (2022): 183-195.