



Intrusion Signalling System by Using AH-MAC in Network-Coded Mobile Small Cells

Chanumolu Kiran Kumar^{*}, Nandhakumar Ramachandran^{*}

SCOPE, VIT-AP University, Amaravati 522241, India

Corresponding Author Email: mounikakiran.138@gmail.com



Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.410423>

ABSTRACT

Received: 31 August 2023

Revised: 20 January 2024

Accepted: 1 April 2024

Available online: 31 August 2024

Keywords:

network coding, small cells, intrusion signal, pollution attack, attacker- location, homomorphic MAC, 5G security

5G networks that can cover urban areas through the use of on-demand, anywhere-and-anytime deployments of small mobile cells made possible by Network Coding (NC) were considered the best schema. Pollution attacks, whereby intermediary nodes alter packets in transit, pose a threat due to their vulnerabilities. The receivers will experience incorrect decoding as a result of these polluted packets. It is critical to identify which packets are polluted in mobile small cells enabled by NC. In a small cell environment enabled by NC, the proposed ISS-AH-MAC (Intrusion Signalling System using Adaptable Homomorphic MAC) may successfully detect polluted packets. Only nodes that have been determined to have high trust levels are allowed to participate in the network's communication. The adaptable variable will be updated only when data packets undergo changes. The attacker and their surrounding areas can be located with relative ease thanks to clustering based on regions. Following identification, it assigns labels to the nodes to help identify and exclude malicious ones during future data transfers. In addition to detecting polluted packets, this method pinpoints the location of the attacker, allowing for the mitigation of future packet pollution to a certain degree. Network intrusion detection is efficient using the suggested approach, which achieves 98% accuracy. Experimental results show that the proposed model achieves better detection accuracy and lower time complexity compared to traditional models.

1. INTRODUCTION

By the year 2025, all of human knowledge will have amassed 175 zeta bytes. Nearly two-thirds had experienced some kind of cybercrime in the last 12 months. Half of those who took the survey think their present security measures aren't enough to stop assaults. The number of security holes that have been reported has been rising steadily. Human mistake is the root cause of 82% of data breaches, according to Verizon's 2022 Data Breach Investigations Report. When it comes to phishing and compromised passwords, human mistake might be a factor. Phishing attacks, in which the attacker attempts to trick the target into giving over personal information or visiting harmful websites, sometimes employ email as a medium. At little over 6 million attacks worldwide in the first half of 2022. Atmost attack bandwidth for attacks like DDoS reached to nearly 999 Gbps, a 57% rise from the second half of 2021, according to Netscout's DDoS Threat Intelligence Report for 2022. The Asia-Pacific region saw a decline of 9% in attacks globally. When looking at North America, the amount of distributed denial of service assaults increased by 2%. Based on these numbers, it's clear that improving Quality of Service requires robust and efficient network intrusion signal models.

5G and subsequent generations are subject to stringent constraints for energy use and data throughput. Increased usage of state-of-the-art smart mobile apps is driving improvements in wireless mobile networking, such as higher

data rates and bandwidth, ultra-low latency, and increased resilience [1]. To meet the requirements of this network, a diverse set of technologies is required. NC has several potential applications, including efficient content distribution, low-latency communication networks, and distributed storage. The system's spectrum efficiency is enhanced since the NC enables storage and material spreading [2].

Wireless communication settings with many paths and hops (MP-MH) are the focus of network coding research. Notwithstanding this, it is essential to deal with every facet of implementing NC. A random linear network coded using a butterfly topology is optimal [3]. Pollution attacks pose a serious threat to the security of circuits that are enabled by NC since intermediate nodes have the ability to recode packets. This study aims to provide hybrid solutions for mobile networks using cryptology. However already some proposals have been mentioned for information theory and cryptographic methods to reduce pollution assaults [4].

To efficiently and affordably provide 5G services to both densely populated urban regions and more remote rural areas, small cell technology is a crucial 5G enabler [5]. Due to power consumption, packet loss, and limited network connectivity, NC technology has been found to be an effective model for increasing throughput in mobile small cells [6]. Despite the possibility of byzantine modification/fabrication assaults caused by network coding, it has been reliable in the past. Instead of storing and transmitting data, NC-enabled technologies allow for its mixing. In addition to XOR and

random linear NC, there are two more techniques to apply NC. When it comes to network coding, homomorphic message authentication helps identify threats better [7].

In the event of a node failure, SDN will keep track of all relevant information and eliminate the affected node from the network [8]. The price for all of these advantages is reduced. However, pollution attacks can happen in NC-enabled wireless networks. This happens when an attacker injects modified incoming packets into the network, making it so that the target nodes can't understand the native packets. One may observe the small cell approach in Figure 1.

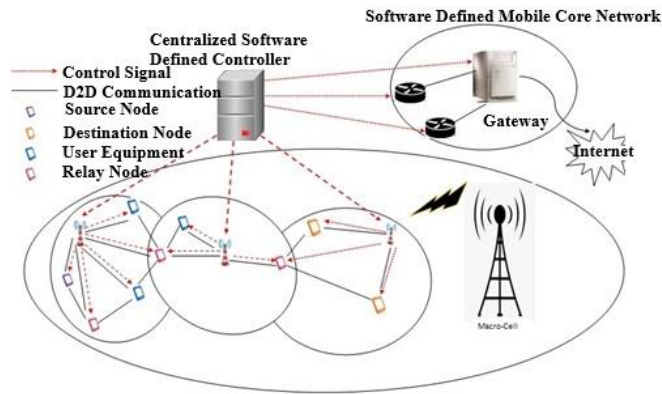


Figure 1. Small cell model

Pollution attacks against privacy, integrity, availability, and authentication can occur in a system powered by NC if hostile intermediary nodes manipulate packets in transit [9]. Possible defenses include cryptography techniques, human intervention, and intrusion signal analysis [10]. These altered packets will lead to incorrect decoding. Through the network, the infected packets propagated, corrupting even more traffic as they evaded actual nodes [11]. Consequently, identifying the harmful users and the polluted packets is of the utmost importance. But most intrusion signal systems on the market today are signature-based, meaning they look for specific characteristics of known attacks in order to identify them. Unable to detect unexpected threats [12], these systems necessitate frequent upgrades to both the rule-based and signatures. Anomaly detection systems, on the other hand, are part of Intrusion Signalling Systems that construct the typical behavior of systems and networks. This makes them highly successful in detecting and thwarting both known and unseen attacks [13]. Anomaly Signalling Systems have great theoretical appeal, but there are several technical hurdles that must be surmounted before they can be extensively used. Among these issues are things like a high false alert rate and an inability to increase up to gigabit speeds, among others [14].

While there are many strategies for avoiding pollution attacks, very few center on tracking down the perpetrators. Any node in a network that deliberately tries to interfere with other nodes' services is said to be malicious [15]. An attack known as packet pollution occurs when bad nodes introduce tainted packets into the network, which include false information [16]. Pollution attacks can compromise packets that be encrypted or not. Data transmission in a network involves a large number of clusters, as seen in Figure 2. From each cluster, a few nodes are chosen to serve as cluster heads, and it is their job to collect data from the rest of the cluster [17]. The network's performance can degrade if the cluster comprises nodes that are malicious or if pollution assaults

spread to other clusters. Figure 2 shows the nodes that are malicious and cause pollution assaults in SC network clusters.

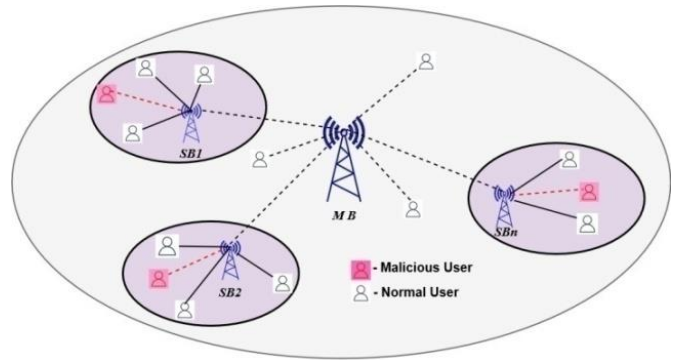


Figure 2. Malicious nodes causing pollution attacks

In congested mobile network situations, optimizing bandwidth requirements and energy usage is made possible via network coding [18]. Security issues persist, however, with network coding installations [19]. A network's SDN controllers can help us detect harmful activities quickly and easily [20]. In addition to enhancing network energy efficiency, implementing NC techniques reduces the quantity of transmissions needed for communications. Dos When it comes to network coding, attacks and pollution attacks are all too common. Homomorphic MAC allows us to efficiently recognize data and label attacks on pollution.

Security studies for mobile devices have occasionally concentrated on physical components and/or methods of access control. But these solutions don't protect private information on lost or stolen devices in the post-authentication phase. Thanks to the capabilities of modern smart devices, users can already collect a mountain of data about their own actions. Also, if your phone is stolen, you may find out exactly where it is because a lot of them have GPS receivers [21]. Few research have been suggested for identifying the perpetrator, despite several defense measures against pollution attacks. This study proposes an intrusion signal system that can detect and stop attacks that cause pollution while also identifying the exact location of the invader. This study recommends an Intrusion Signalling System by using Adaptable Homomorphic MAC (ISS-AH-MAC) to properly detect intrusions in small cells that have NC capabilities. While data is being transmitted, it is checked at each node to make sure no malicious operations, like adding fake data to the original data, are happening. If so, the model suggests using region-based networks to cluster nodes. When applied to the system, the proposed model improves performance by reliably detecting intrusions.

1.1 Small cells enabled by network coding and their vulnerabilities

Passwords, packets and keys are sensitive information that an eavesdropper can gain by tapping into a wireless transmission or one or more cable links. Regarding this issue, packets are neither altered or injected by eavesdroppers. All they have to do is get the important, private data that has to be transmitted by observing the links. Therefore, it is considered a passive attack to listen in on someone's talk. Eavesdropping is possible by both benign and malicious intermediate nodes. If they have access to enough linearly distinct permutations of

packets, they can decode them and get all the information.

Analyzing network traffic is a common tactic for hackers targeting wireless networks. The attackers behind a traffic analysis assault can learn the ins and outs of the network's design and the packets' destinations by monitoring all communications [22]. By examining and monitoring network data, malevolent actors jeopardize network confidentiality. There is a risk to all network coding protocols, both state aware and stateless ones.

When more than 2 malicious nodes compromised to construct a subway that links two additional nodes, it is known as a wormhole attack. Subsequently, they persuade the adjacent nodes that the tunnel's two termini are equally distant. Data packets can be captured and sent back via the tunnel by wormhole intruders [23]. The potential ramifications of wormhole attacks are greater for state-aware NC protocols than for stateless protocols.

In an entropy attack, the perpetrator transmits data packets that the target systems are already aware of. More specifically, the bad node generates a new coded packet which isn't particularly creative; Due to this, the delivery of packets is sequentially reliant on the node downstream's storage of coded packets. This is fine, but rigidly sequentially coded packet is useless since the receivers get no data that might aid in deciphering the original packets.

2. LITERATURE SURVEY

Parsamehr et al. [1] offered an Intrusion detection and prevention method for Mobile small cells enabled by NC for the first time. When an intrusive event is detected, homomorphic message authentication codes (MACs) enhanced in null space are employed for detect pollution assaults and restrict risk. In order to prevent network intrusions, the author planned to do future research into expanding the proposed IDPS into a collaborative IDPS. In this way, polluting assaults may be pinpointed and their origins located, making it much easier to stop intruders from accessing the network. Additionally, this research aims for enhancing the proposed model in a way, so, it might also aid in fixing corrupted packets once they were found.

Attackers may still manage to tamper with packets during future source-to-destination coded packet communication, regardless of how much effort is put into identifying pollution attacks. In order to stop additional data packet pollution, Parsamehr et al. [2] presented a narrative method for detection and prevention of untrusions (IDL) that finds the attacker's precise location and drops polluted packets. IDL detects and locates using a homomorphic MAC approach based on null space. An efficient method to prevent Mobile small cells enabled by NC from resource drain is IDL, which does not require initial use on all mobile devices. It is necessary to decrease the proposed model's significant network overhead in order to improve the network's performance.

Congestion and restrictions on radio channels can lead to a wide range of retransmission requests. The model can use the topology knowledge of the network to find the best-encoded packets that have gone missing. The global communication topology of the proposed TANC system can make use of software-defined networking controllers. Regular analysis of the nodes can improve speed and prevent the topology-aware model's increased computing cost. Neighbor node accessing models can also avoid this.

SECRET, a Training network created by Rodriguez et al. [4], provides a great setting for ESRs to learn about 5G wireless communication systems. The European Commission has authorized the project's financing via the Marie Curie People Program, which is a component of the H2020 research and innovation program. By training and teaching 17 ESRs, this project hopes to bridge the gap between the capabilities of today's networking tools and those needed to meet anticipated demands in 2020. The new "femto cell" type of cells that SECRET is deploying are called "mobile small cells," and they offer reduced costs per bit, higher capacity, and the ability to handle more customers. Due to the inclusion of complicated procedures, the suggested model has a high time complexity. Lightweight procedures can be used to process femto cells, which improves the network's performance.

The new paradigm has the potential to improve mobile networks in particular by reducing Delivery of packets in multicast network, as described by Parsamehr et al. [5], which in turn increases network capacity, becomes more resilient to packet losses, and uses less energy. Regardless of these significant benefits, mobile small cells allowed by NC are susceptible to many attacks due to the intrinsic weaknesses of NC. This study classifies potential weak spots in small cell security that are facilitated by NC.

The demand for faster data speeds is skyrocketing, and 5G networks might take advantage of this by deploying network coding-enabled mobile small cells on any device, at any time. Although there are many benefits to 5G mobile small cell network coding technology, it cannot reach its full potential until pollution attacks are addressed. An IDPS system was suggested by Parsamehr et al. [6]. In addition to identifying and preventing pollution attacks, it can pinpoint the exact location of the malicious nodes responsible for the problem.

Wireless networks can improve their throughput, energy consumption ability to bounce back to packet failures, delay, and energy consumption with the help of NC technology, according to a new study by Esfahani et al. [7]. An extremely dangerous security threat for NC-enabled wireless networks is data pollution attacks, in which an adversarial node can introduce polluted packets into the system, making it impossible for the intended nodes to decode them.

Making sure wireless networks are resilient and dependable is crucial if we wish to meet the demands of future generations of networks. For high-density mobile networks, network coding is a key enabler for better energy consumption and bandwidth requirements. Before coding approaches may be used for future mobile network installations, security considerations must be addressed. It is possible to construct very efficient and readily flexible network topologies in the context of mobile small cells by employing software defined networking. Small cells with secure network coding with little delay were studied by Adat et al. [8] using SDN-based mobile small cells.

The use of a network coding scheme, as suggested by Adat et al. [9], could shield enhanced blockchain with mobile small cell environment against pollution attacks. Along with the architecture of the little SECRET blockchain, we give an introduction to overhead transmission and latency issues. Small mobile cells can significantly improve network performance as a whole thanks to network coding. It is of the utmost importance to shield little SECRET cells from pollution. The security of the system is compromised because attackers can easily crack the conceptual model cryptography techniques. The suggested architecture can be enhanced with

the use of digital signatures, hashing, and encryption techniques.

3. PROPOSED MODEL

When it comes to network code, the two most dangerous threats are pollution and modification attacks. By inserting malicious packets, an attacker aims to exhaust the network's resources [24]. Nodes generate coded packets that pollute the network significantly when they combine dirty packets with clean ones. Consequently, the native packets cannot be decoded by the receivers. Their ability to pollute different environments allows us to classify them into two broad groups.

- Data Pollution
- Tags Pollution

In a data pollution assault, the enemy tampers with data packets. An attacker can compromise a network's data decoding capabilities by manipulating the tags attached to packets. An outside threat actor or a compromised node within the network could potentially launch a pollution assault. Using Mobile small cells enabled by NC, the suggested model can detect pollution attacks, pinpoint their exact location, and attempt to prevent such attacks in subsequent transmission rounds. The ISS-AH-MAC method, which is optimized for mobile small cells and utilized for both detection and localization, is based on adaptive variables and homomorphic messages. The detecting mechanism presented in ISS-AH-MAC is used to intercept, after discard malicious (polluting) packets at the first stab. It is possible for attackers to corrupt messages even in travel from the to their intended recipients from source node leading to bandwidth waste in the network. We need to find out where the attackers are exactly so we can stop them from spreading farther. Figure 3 depicts the model framework that has been suggested.

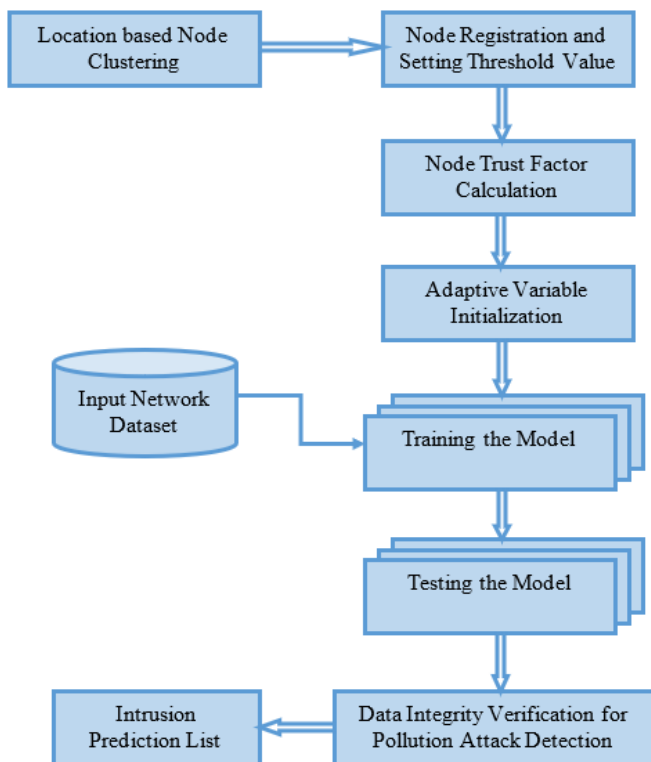


Figure 3. Proposed model framework

Clustering nodes according to how similar their base addresses are is the first step in the suggested model. The regional administrator (RA) is chosen based on the node's accessibility to nodes in that cluster region. The regional administrator (RA) is responsible for maintaining the information and registering nodes from the created regions (clusters). To establish the starting points for the threshold values, nodes analyze the packet delivery rate (PDR) of each transmission to their neighbors and start the transmission process among their regional members [25]. In order to keep their regional administrator (RA) up-to-date, all nodes must keep track of their neighbors' PDR history. This history makes it easy to see how those numbers have changed over time. energy consumption, PDR, data packets dropped, and ability to compute are the factors that will be used to calculate the trust factor of the registered nodes [26]. The trust factor is useful for distinguishing between benign and malevolent nodes. Nodes report back to their RA's; after certain intervals of iteration [27]. If a node's performance deviates from its previously recorded values the neighboring nodes will promptly notify the regional administrator (RA). The regional administrator (RA) then decides whether to accept the node in the region or suspend it after reviewing its history data and any concerns brought up by neighboring nodes. The node will remain in a suspended status for the time being [28]. The nodes limit initiates the adaptive variable of the proposed model. Both detecting a security breach and pinpointing the location of a malevolent user are equally important for notifying other nodes [26].

Because an adversarial node will fool the network This situation can lead to network instability, if an adversarial node tricks the network into thinking a valid parent node is an enemy [29]. So, more methods for finding or validating attackers are required to keep a fair communication network running. Only nodes that have been determined to have high trust levels are allowed to participate in the network's communication [30]. Whenever an alteration in the data packets, the value of the adaptable variable will be updated. For Intrusion Signalling System based on H-MAC, this research proposes an ISS-AH-MAC.

Using the Homomorphic MAC technique, network pollution attacks can be identified. The H-MAC generates a new set of keys for every encoding and decoding process. Additionally, the key production mechanism makes it harder for attackers to regenerate keys for reuse [31]. The HMAC algorithm is a method for authenticating messages that uses a cryptographic key in conjunction with a hash function. Through the use of HMAC, unique private keys are assigned to both the server and the client. Possible eavesdroppers can alter the data during transmission of communications. By mixing secret keys during transmission and hashing messages, HMAC helps to prevent this [32]. Once the receiver has received the communication, they can verify its authenticity by utilizing the identical secret key that was employed for decryption. If the receiver recalculates the hash using the same secret key, they will be able to detect any changes done by the attackers.

The proposed model takes into consideration all transaction data included in the network dataset. The dataset is partitioned into two sets, one for training and one for testing, with a ratio of 80:20. Details regarding a significant number of network attacks are included in the network dataset under examination, which is used to execute the training operation. The training is used to test the network samples. After pre-processing, the

dataset is free of null values, making it clean. For the purpose of enhancing service quality, each node in the network has been assigned with a different adaptive variable. Any modification to this variable will trigger incursions, reveal the nodes' locations, and ultimately lead to their removal from the network. The adaptive variable is validated on a regular basis. Changing an adaptable variable on a node can cause an incursion. The malicious nodes will be removed from the network as a preventative measure.

Users should consider the organization's unique information security requirements, integrate with the existing IT infrastructure, and minimize disruption when selecting an intrusion detection system (IDS) and planning for its installation and rollout. The ultimate goal of intrusion detection systems is to detect anomalies, not cause them. It is crucial to consider the level of experience of the individuals tasked with system installation, configuration, and maintenance. For acquiring required authorization and authentication for its number of remote agents, the complex IDS must be developed. As soon as the IDS is live, all of its components must be configured, fine-tuned, tested, and monitored. In order to accurately detect pollution attacks, this suggested model must be implemented in the telecommunications industry. It does this by detecting changes in adaptive variables for each data packet type. Accurately detecting pollution assaults and the nodes causing them can improve network performance and, by extension, service quality.

When compared to the more traditional store-and-forward strategy, data pollution attacks targeting NC have a better chance of success. The capacity of the forwarders to identify data pollution assaults is crucial for the sink nodes to correctly retrieve source messages. To emphasize the point, even a small number of infected messages can infect numerous nodes downstream due to the fact that pollution spreads via recoding. Tag pollution attacks are more sophisticated forms of pollution assaults that have the ability to impact networks that are enabled by NC. The goal of a tag pollution assault is to change the tags that communications carry rather than the contents themselves, which is different from a content-based attack. Even after passing through multiple checkpoints, a message with compromised tags may still be denied. However, this still results in the network's bandwidth being squandered.

Any component of a system, be it software, hardware, or administrative processes, could have a security hole in its network. Threats to networks might be either physical or non-physical. Data and software vulnerabilities are examples of non-physical types of weaknesses. Hackers will have complete control of the network if the IT department does not fix vulnerable operating systems. Intentionally installing malicious software on a computer can lead to a system-wide infection. The foundation of any secure network is its capacity to maintain user privacy, data integrity, and service availability. This concept proposes a homomorphic MAC as a means of user authentication and attack detection in polluted networks.

When it comes to designing a private network, the proposed model is useful for large organizations. By adopting it, they can accurately detect intrusions that could cause data loss, improve data delivery rate, and maintain security. This, in turn, helps to prevent attacks from spreading to other branches of the organization. With the advent of 5G network services, this model aids the telecom industry in improving their model for intrusion signal detection which leads to improving the

Quality in the service. The following algorithm will discuss the steps in the proposed model.

Input: Network Dataset.

Output: Malicious Nodes causing pollution.

A: Nodes Clustering.

B: Registration of Nodes.

C: Threshold Calculation.

D: Determine the trusts' factors.

E: Generating Homomorphic-MAC Key Pairs.

F: Initializing the Adaptive Variables.

G: Attack Detection for Pollution.

H: Create a Predicted List.

3.1 Node clustering

Clustering, another name for cluster analysis, is dividing the network nodes into smaller groups according to the degree to which their individual parts are similar. Using the small cells dataset as a starting point, the suggested model clusters nodes according to their locations. Using the suggested clustering paradigm, the entire network is partitioned into numerous subnetworks.

Node clustering is conducted by taking into account the geographical location of each node, with respect to the base addresses of nodes registered within the network.

$$Nodeinf[M] \leftarrow getnodeID(Nid) \in SCD \quad (1)$$

$$Simm(NodeID(i)) \leftarrow \sum_{i=1}^M diff(phyaddr(Nid(i)), phyaddr(Nid(i+1))) \quad (2)$$

return 1 if true
return 0 Otherwise

$$CNset[N] \leftarrow \sum_{i=1}^N \max(Simm(Nid(i))) \in Nodeinf \quad (3)$$

getnodeID() considers each node ID. The central network administrator assigns each node a unique identification. The location identity of each node is represented by its baseaddr. All of the nodes will be grouped in to a cluster, with nodes from the same address being in the same set, depending on the difference in their base addresses. The list of network nodes must contain all of the nodes in the cluster set. Any malicious node that joins the network without an identifier will be added to the cluster.

3.2 Registration and trust factor calculation of nodes

Network service provider will keep track of the incoming nodes so they may be identified in the future when data is transmitted. The procedure for registering nodes is carried out as:

$$Nidinf \leftarrow Node(phyaddr) \in SCD \quad (4)$$

$$TI(Nid) \leftarrow time(HHMMSS) \quad (5)$$

$$r \leftarrow rand(Th, Nid) \quad (6)$$

$$NUniqID(i) \leftarrow r * Nid + TI \quad (7)$$

$$Th = \sum_{i=1}^N \sqrt{\sum_{i=1}^M \frac{\min(NID) + \max(NID)}{\text{count}(\text{phyaddr}(NID))}} \quad (8)$$

The base address of the node is the most important parameter in the suggested model's node registration procedure. To further filter out fraudulent nodes, we take into account the moment a node joins the network. During node registration, the random() function is used to consider a once-used value that aids in identifying malicious nodes that masquerade as normal nodes and either allows them to access the network or prevents them from doing so.

Deploying a trust management system for WSNs can support the network's decision-making procedures [11]. When people in a WSN use it, they are better able to handle the uncertainty of not knowing what their peers will do next. There are many various types of security requirements that must be satisfied due to the vast array of purposes for WSNs. Therefore, for a WSN to last, its nodes need to be reliable and work together. Therefore, nodes must have the ability to trust each other. Assuming node malice is absent, a trust factor will be allocated to each node in the cluster. This is how the trust factor is determined:

$$TrFact(Nid) = \frac{\lambda - \text{availene}(NodeID(i)) + \max(PDR) + \sum_{i=1}^m \mu(NodeID(i))}{\begin{cases} ITr \leftarrow \text{rand}(\lambda, \mu) > 50 \text{ if } V > Th \\ ITr \leftarrow \text{rand}(1, 50) \end{cases}} \quad (9)$$

In this case, λ represents the network's maximum energy allocation and μ stands for the node's load in the cluster. Th stands for the created value, V is threshold value. The trust factor will be assigned by using threshold value. The secret value from a service provider is the threshold value, which is utilized to construct the unique label for each node.

3.3 Homomorphic MAC calculation

Network pollution attacks can be detected using the Homomorphic MAC procedure. The H-MAC creates a set of keys that can be used for encoding and decoding just once. It is difficult for attackers to reproduce keys for reuse because of this key creation procedure. The procedure for calculating the H-MAC is carried out as:

a (i, j, M) is a homomorphic MAC.

To generate this a random calculator was used.

$$R \leftarrow (Key_{Nid} + Th) \leftarrow \text{rand}(TrFact, NUniqID) \quad (10)$$

The H-MAC model uses the keys it generates to compute a key pair, which is then utilized to evade pollution tags.

$$\sum_{i=1}^m TrFact(NID(i)) + R * Th \ll 2 \rightarrow Key_i \quad (11)$$

$$\sum_{j=1}^m R - TrFact(NID(j+1)) + Th \ll 4 - \mu \rightarrow Key_j \quad (12)$$

$$\text{H-MAC } \{\text{Keyset } [M]\} \rightarrow \sum_{i,j=1}^M \{Key_i : Key_j\} \quad (13)$$

A polluting node with TrFactor(i) that views packet 'p' as any pollution packet, such as a(p), will be revealed if wrong value of a Node Unique ID(i) is detected.

3.4 Adaptive variable for detection

Whenever there is a change in the data bits produced by a pollution attack, the network will update every trusted node's adaptive variable. When the adaptive variable changes, a pollution attack happens, and if the node's location is known, it can be removed from the network to keep it secure. Every node receives the adaptive variable, and pattern analysis is performed at each node. By observing how the adaptive variable changes at a certain node, we can pinpoint when the pattern has changed. When a node's adaptive variable is altered, it is considered malicious and removed from the network.

Every node in the network has an adaptive variable set up from beginning so that any changes to the value can be used to detect attacks.

$$AV = \max(PDR) * R + Key_i \quad (14)$$

$$AV(NUniqID(i)) = \sum_{j=1}^M \mu - \text{loss}(p) + R + Key_j \quad (15)$$

In order to detect a network pollution attack, the nodes might send data to other clusters and confirm that the data is intact. The process of identifying an attack on pollution is carried out as:

$$\begin{aligned} &Polluattak[CNset(AV)] = \\ &\sum_{i=1}^M \max(PDR) - P(\text{Mesg}) + \\ &\frac{(\text{Min}(Cset(i)) - TrFact(NID)) - (\text{sim}(P, P+1))}{(\text{sizeof}(CNset(i)))} \quad (16) \\ &\begin{cases} V \leftarrow 0 \text{ if } \text{sim}(AV') \in AV \\ V \leftarrow 1 \text{ if } \text{sim}(AV') \notin AV \end{cases} \end{aligned}$$

Here Mesg is the packet message, P is the packet considered and P+1 is the next packet and V is the value generated.

Any time the adaptive variable changes, an intrusion prediction is created. A list of predictions is produced by:

$$\begin{aligned} &PattackLis(Cset(i)) \leftarrow \\ &\left(\sum_{i=1}^N \left(\sum_{j=1}^M \frac{\mu TrFact(j)}{R \lambda} + \text{sim}(AV, AV') \right) + \right. \\ &\left. \text{diff}(P(\text{Mesg}), P+1(\text{Mesg})) \right) \\ &\leftarrow NUniqID(i) \quad (17) \end{aligned}$$

4. RESULTS

In an NC-enabled system, packets can be corrupted while in transit due to pollution assaults. Changing packets create decoding issues. Permitting infected packets to pass through valid nodes increases the chances of other packets to infect. So, identifying the infected packets and identifying the bad actors are of equal importance. With so many safeguards in place to prevent pollution attacks, it is puzzling that so few integrity

schemes prioritize identifying malicious individuals. To test the suggested model in NS2 Simulator and Python for pollution attack detection, we used the TCL scripting language. We also ran the model in Google Co-lab. You can find the public dataset service provider's benchmark dataset at this URL. <https://www.kaggle.com/datasets/ymirsky/network-attack-dataset-kitsune>.

The proposed model exclusively focuses on pollution attacks, although nine datasets detailing attack patterns have been gathered from either an IP-based commercial monitoring system or a network featuring Internet of Things devices. Each dataset contains millions of network packets from different cyber assaults, which researchers can use to study and detect attacks using the dataset parameters. our proposed ISS-AH-MAC is compared with the traditional IDLAPM-NCEMSC Model [2], CDIDRC [11], CSM [12] and D2Gen Model [13]. The results show that the suggested model outperforms the existing models when it comes to detecting threats related to pollution.

Registration nodes allow mobile nodes to notify their base agents of their present coverage status. The registration procedure aids in the proper identification of nodes, and once registered, mobile small cell nodes are able to perform a multitude of roles. In Table 1 and Figure 4, we can see the levels of node registration time for both the existing and proposed models.

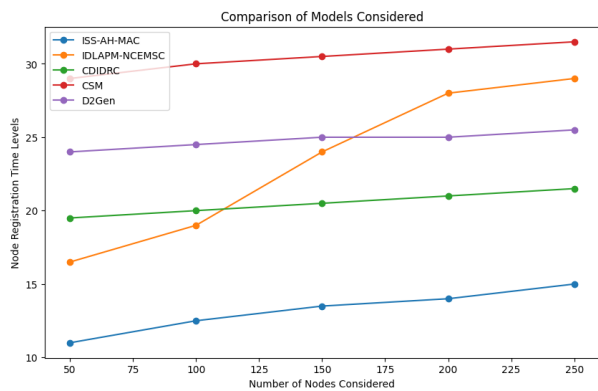


Figure 4. Node registration time levels

Table 1. Node registration time levels

Number of Nodes Considered	Models Considered				
	Proposed ISS-AH-MAC Model	IDLAPM-NCEMSC Model	CDIDRC Model	CSM Model	D2Gen Model
50	11	16.5	19.5	29	24
100	12.5	19	20	30	24.5
150	13.5	24	20.5	30.5	25
200	14	28	21	31	25
250	15	29	21.5	31.5	25.5

Table 2. Location based node clustering time levels

Number of Nodes Considered	Models Considered				
	Proposed ISS-AH-MAC Model	IDLAPM-NCEMSC Model	CDIDRC Model	CSM Model	D2Gen Model
50	22	38	30	42.5	47.5
100	23	39	30.5	42.5	48
150	24	40	30.5	43	49
200	25	40	31	44	50
250	26	41	32	45	51

Clustering entails assembling a collection of nodes from close proximity or shared characteristics into a single entity. The clustering technique relies heavily on the nodes' addresses. The position in a network is the determining factor in the identity of its nodes. It is not feasible to use a 64-bit address structure to accommodate the vast number of unique IDs due to the dense deployment of sensor nodes. As part of this paradigm, the cluster's nodes exchange data using their local IP addresses. Table 2 and Figure 5 illustrate the time levels, and the suggested model's location-based node clustering is faster than the conventional ones.

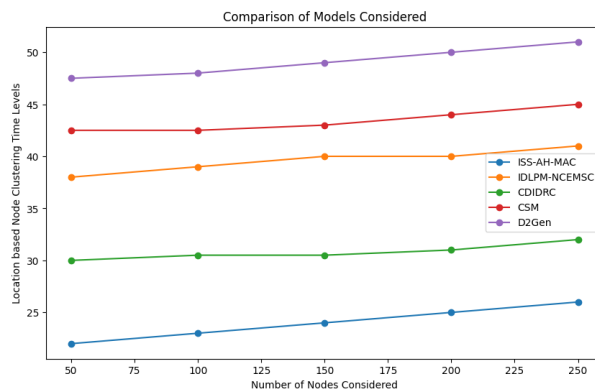


Figure 5. Location based node clustering time levels

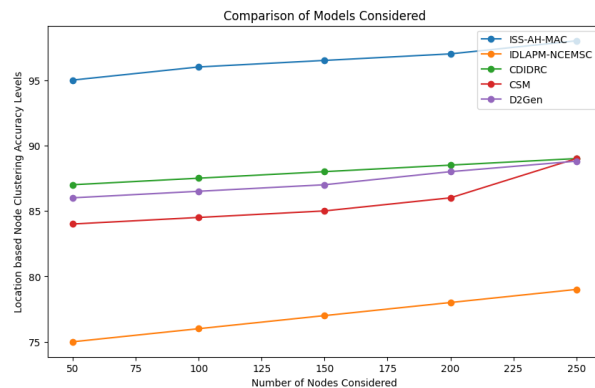


Figure 6. Location based node clustering accuracy levels

Table 3. Location based node clustering accuracy levels

Number of Nodes Considered	Models Considered				
	Proposed ISS-AH-MAC Model	IDLAPM-NCEMSC Model	CDIDRC Model	CSM Model	D2Gen Model
50	95	75	87	84	86
100	96	76	87.5	84.5	86.5
150	96.5	77	88	85	87
200	97	78	88.5	86	88
250	98	79	89	89	88.8

Table 4. Trust factor generation accuracy levels

Node Registration Time Levels	Models Considered				
	Proposed ISS-AH-MAC Model	IDLAPM-NCEMSC Model	CDIDRC Model	CSM Model	D2Gen Model
50	96	62	90	81	77
100	96	62	90.5	83	77.5
150	97	63	91	87	78
200	97	64	91	89	78
250	98	65	91.5	90	78

In order to facilitate the analysis of node behavior, location-based node clustering groups nodes that are in close proximity to one another or that have comparable physical locations. Both the current and proposed models' accuracy levels in location-based node clustering are displayed in Table 3 and Figure 6, respectively.

Based on its history of data transmissions, a node's trust factor indicates whether it is a benign node or one with malicious characteristics. To prevent assaults on the network, only trusted nodes are considered for data transfer based on the node's trust factor. Table 4 and Figure 7 display the levels of accuracy for trust factor creation.

Data or information stored on a network is considered to have integrity if it is secure from tampering or deletion by unauthorised users. This kind of data management is crucial for ensuring that data is clean, reliable, and accurate. The time levels for data integrity verification are displayed in Table 5 and Figure 8.

Pollution assault occurs when malicious nodes introduce corrupted data packets into the network or corrupt current packets. In a network, packets, whether plain or encoded, can become corrupted. The degrees of accuracy for detecting pollution attacks are displayed in Table 6 and Figure 9.

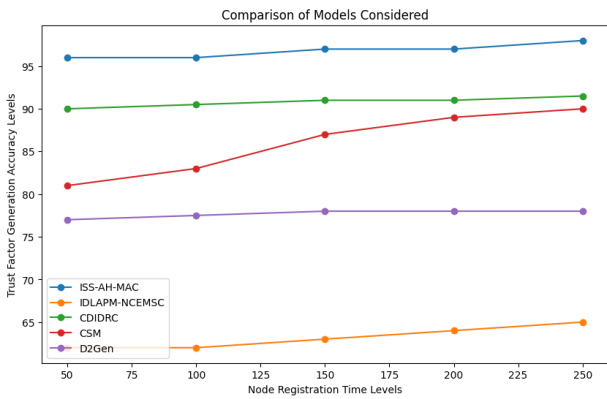


Figure 7. Trust factor generation accuracy levels

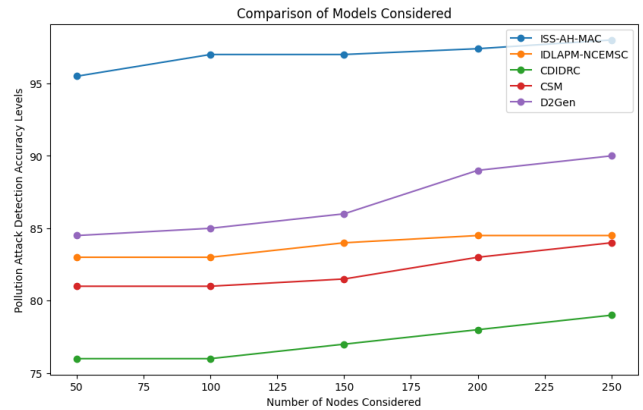


Figure 9. Pollution attack detection accuracy levels

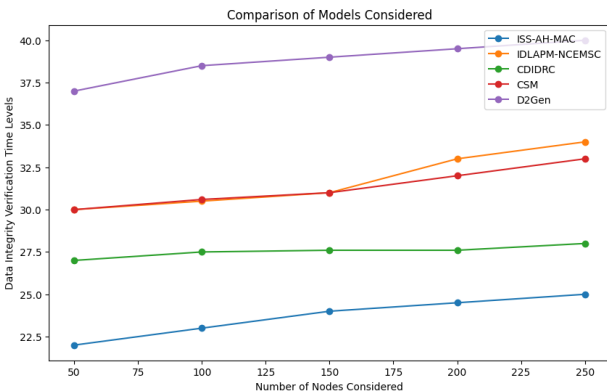


Figure 8. Data integrity verification time levels

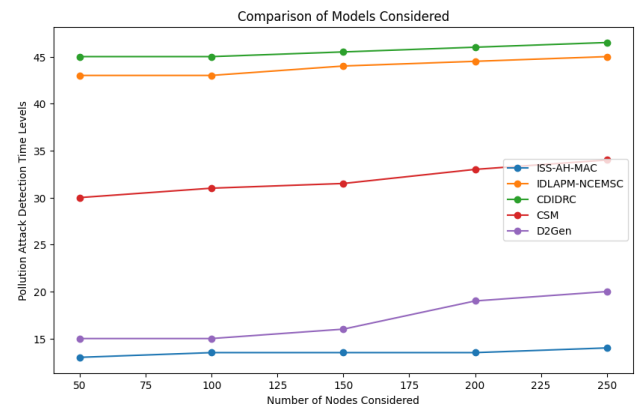


Figure 10. Pollution attack detection time levels

Table 5. Data integrity verification time levels

Number of Nodes Considered	Models Considered				
	Proposed ISS-AH-MAC Model	IDLAPM-NCEMSC Model	CDIDRC Model	CSM Model	D2Gen Model
50	22	30	27	30	37
100	23	30.5	27.5	30.6	38.5
150	24	31	27.6	31	39
200	24.5	33	27.6	32	39.5
250	25	34	28	33	40

Table 6. Pollution attack detection accuracy levels

Number of Nodes Considered	Models Considered				
	Proposed ISS-AH-MAC Model	IDLAPM-NCEMSC Model	CDIDRC Model	CSM Model	D2Gen Model
50	95.5	83	76	81	84.5
100	97	83	76	81	85
150	97	84	77	81.5	86
200	97.4	84.5	78	83	89
250	98	84.5	79	84	90

Table 7. Pollution attack detection time levels

Number of Nodes Considered	Models Considered				
	Proposed ISS-AH-MAC Model	IDLAPM-NCEMSC Model	CDIDRC Model	CSM Model	D2Gen Model
50	13	43	45	30	15
100	13.5	43	45	31	15
150	13.5	44	45.5	31.5	16
200	13.5	44.5	46	33	19
250	14	45	46.5	34	20

In particular, networks that employ network coding are vulnerable to pollution and other targeted attacks. A malicious user can inject a corrupted packet into the transition since intermediate nodes are permitted to recode packets. Threats to network security might arise when corrupted packets are transmitted alongside valid ones. The throughput of a network is significantly reduced in polluting assaults. Table 7 and Figure 10 display the levels of time it takes for the proposed model and existing methods to detect pollution attacks in the network, respectively.

5. CONCLUSIONS

Network coding can improve network resilience and capacity, as has been demonstrated. Data integrity cannot be confirmed using normal MACs and checksums due to packet alteration by intermediate nodes. In networks that use encoding, it only takes one bad node to overwhelm the system with incorrect packets, making it impossible for the recipient to decode them correctly. The ultra-reliable, high-throughput requirements of 5G could be met by mobile small cells that have network coding capabilities. However, safeguarding the system from external and internal dangers is essential for achieving its full potential. Most often, an infected node will launch a pollution attack, flooding the network with fake packets that the receiver will not be able to decode. In order to reliably identify network intrusion signals, this study introduces an adaptive variable based pollution attack detection model. It is necessary to use different integrity methods than what is currently in use because network coding permits packets to be mixed together in real time. Detecting network intrusion signals is a breeze using the suggested

approach, which hits 98% accuracy. Future work on improving detection accuracy can take into account feedback and node loss rate.

REFERENCES

- [1] Parsamehr, R., Esfahani, A., Mantas, G., Radwan, A., Mumtaz, S., Rodriguez, J., Martínez-Ortega, J.F. (2019). A novel intrusion detection and prevention scheme for network coding-enabled mobile small cells. *IEEE Transactions on Computational Social Systems*, 6(6): 1467-1477. <https://doi.org/10.1109/TCSS.2019.2949153>
- [2] Parsamehr, R., Mantas, G., Rodriguez, J., Martínez-Ortega, J.F. (2020). IDLP: An efficient intrusion detection and location-aware prevention mechanism for network coding-enabled mobile small cells. *IEEE Access*, 8: 43863-43875. <https://doi.org/10.1109/ACCESS.2020.2977428>
- [3] Chen, Y.J., Wang, L.C., Wang, K., Ho, W.L. (2018). Topology-aware network coding for wireless multicast. *IEEE Systems Journal*, 12(4): 3683-3692. <http://doi.org/10.1109/JSYST.2018.2802493>
- [4] Rodriguez, J., Radwan, A., Barbosa, C., Fitzek, F.H., Abd-Alhameed, R.A., Noras, J.M., Jones, S.M.R., Politis, I., Galiotos, P., Schulte, G., Rayit, A., Sousa, M., Alheiro, R., Gelabert, X., Koudouridis, G.P. (2017). SECRET—Secure network coding for reduced energy next generation mobile small cells: A European Training Network in wireless communications and networking for 5G. In *2017 Internet Technologies and Applications (ITA)*, Wrexham, UK, pp. 329-333. <http://doi.org/10.1109/ITECHA.2017.8101964>

- [5] Parsamehr, R., Mantas, G., Radwan, A., Rodriguez, J., Martinez, J.F. (2019). Security threats in network coding-enabled mobile small cells. In *Broadband Communications, Networks, and Systems: 9th International EAI Conference, Broadnets 2018, Faro, Portugal*, pp. 337-346. https://doi.org/10.1007/978-3-030-05195-2_33
- [6] Parsamehr, R., Esfahani, A., Mantas, G., Rodriguez, J., Martinez-Ortega, J.F. (2019). A location-aware IDPS scheme for network coding-enabled mobile small cells. In *2019 IEEE 2nd 5G World Forum (5GWF), Dresden, Germany*, pp. 91-96. <https://doi.org/10.1109/5GWF.2019.8911650>
- [7] Esfahani, A., Mantas, G., Rodriguez, J., Neves, J.C. (2017). An efficient homomorphic MAC-based scheme against data and tag pollution attacks in network coding-enabled wireless networks. *International Journal of Information Security*, 16: 627-639. <https://doi.org/10.1007/s10207-016-0351-z>
- [8] Adat, V., Politis, I., Tselios, C., Kotsopoulos, S. (2018). Secure network coding for SDN-based mobile small cells. In: Sucasas, V., Mantas, G., Althunibat, S. (eds) *Broadband Communications, Networks, and Systems. BROADNETS 2018. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Springer, Cham. https://doi.org/10.1007/978-3-030-05195-2_34
- [9] Adat, V., Politis, I., Tselios, C., Galiotos, P., Kotsopoulos, S. (2018). On blockchain enhanced secure network coding for 5G deployments. In *2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates*, pp. 1-7. <https://doi.org/10.1109/GLOCOM.2018.8647581>
- [10] Hansen, J., Krigslund, J., Lucani, D.E., Pahlevani, P., Fitzek, F.H. (2018). Bridging inter-flow and intra-flow network coding in wireless mesh networks: From theory to implementation. *Computer Networks*, 145: 1-12. <https://doi.org/10.1016/j.comnet.2018.07.014>
- [11] Segura, G.A.N., Chorti, A., Margi, C.B. (2021). Centralized and distributed intrusion detection for resource-constrained wireless SDN networks. *IEEE Internet of Things Journal*, 9(10): 7746-7758. <https://doi.org/10.1109/IJOT.2021.3114270>
- [12] Raouf, A., Lung, C.H., Matrawy, A. (2021). Securing RPL using network coding: The chained secure mode (CSM). *IEEE Internet of Things Journal*, 9(7): 4888-4898. <https://doi.org/10.1109/IJOT.2021.3109109>
- [13] Makhdoom, I., Hayawi, K., Kaosar, M., Mathew, S.S., Ho, P.H. (2021). D2Gen: A decentralized device genome based integrity verification mechanism for collaborative intrusion detection systems. *IEEE Access*, 9: 137260-137280. <https://doi.org/10.1109/ACCESS.2021.3117938>
- [14] Esfahani, A., Mantas, G., Silva, H., Rodriguez, J., Neves, J.C. (2016). An efficient MAC-based scheme against pollution attacks in XOR network coding-enabled WBANs for remote patient monitoring systems. *EURASIP Journal on Wireless Communications and Networking*, 2016: 1-10. <https://doi.org/10.1186/s13638-016-0601-9>
- [15] Esfahani, A., Mantas, G., Rodriguez, J., Neves, J.C. (2017). An efficient homomorphic MAC-based scheme against data and tag pollution attacks in network coding-enabled wireless networks. *International Journal of Information Security*, 16: 627-639. <https://doi.org/10.1007/s10207-016-0351-z>
- [16] Ferrag, M.A., Maglaras, L., Argyriou, A., Kosmanos, D., Janicke, H. (2018). Security for 4G and 5G cellular networks: A survey of existing authentication and privacy-preserving schemes. *Journal of Network and Computer Applications*, 101: 55-82. <https://doi.org/10.1016/j.jnca.2017.10.017>
- [17] Patcha, A., Park, J.M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks*, 51(12): 3448-3470. <https://doi.org/10.1016/j.comnet.2007.02.001>
- [18] Saini, R., Khari, M. (2011). Defining malicious behavior of a node and its defensive methods in ad hoc network. *International Journal of Computer Applications*, 20(4): 18-21. <https://doi.org/10.5120/2422-3251>
- [19] Yao, H., Silva, D., Jaggi, S., Langberg, M. (2014). Network codes resilient to jamming and eavesdropping. *IEEE/ACM Transactions on Networking*, 22(6): 1978-1987. <https://doi.org/10.1109/TNET.2013.2294254>
- [20] Fiandrotti, A., Gaeta, R., Grangetto, M. (2018). Securing network coding architectures against pollution attacks with band codes. *IEEE Transactions on Information Forensics and Security*, 14(3): 730-742. <https://doi.org/10.1109/TIFS.2018.2859583>
- [21] Esfahani, A., Mantas, G., Yang, D., Nascimento, A., Rodriguez, J., Neves, J. (2015). Towards secure network coding-enabled wireless sensor networks in cyber-physical systems. *Cyber-Physical Systems: From Theory to Practice*, 395-414.
- [22] Dhakne, A.R., Chatur, P.N. (2016). TCNPR: Trust calculation based on nodes properties and recommendations for intrusion detection in wireless sensor network. *International Journal of Computer Science and Network Security (IJCSNS)*, 16(12): 1.
- [23] Pahlevani, P., Khamfroush, H., Lucani, D.E., Pedersen, M.V., Fitzek, F.H. (2016). Network coding for hop-by-hop communication enhancement in multi-hop networks. *Computer Networks*, 105: 138-149. <https://doi.org/10.1016/j.comnet.2016.05.012>
- [24] Krigslund, J., Hansen, J., Lucani, D.E., Fitzek, F.H., Médard, M. (2015). Network coded software defined networking: Design and implementation. In *Proceedings of European Wireless 2015, 21th European Wireless Conference*, pp. 1-6.
- [25] Rajeshkumar, G., Valluvan, K.R. (2017). An energy aware trust based intrusion detection system with adaptive acknowledgement for wireless sensor network. *Wireless Personal Communications*, 94: 1993-2007. <https://doi.org/10.1007/s11277-016-3349-y>
- [26] Bakhsh, S.T., Alghamdi, S., Alsemmeiri, R.A., Hassan, S.R. (2019). An adaptive intrusion detection and prevention system for Internet of Things. *International Journal of Distributed Sensor Networks*, 15(11): 1550147719888109. <https://doi.org/10.1177/1550147719888109>
- [27] Jiang, W., Li, Z., Tan, K., Guan, Y., Tong, W. (2021). An adaptive intrusion detection algorithm for in-vehicle CAN bus based on periodicity of message. In *Journal of Physics: Conference Series*, 1748(3): 032023. <https://doi.org/10.1088/1742-6596/1748/3/032023>
- [28] Teng, S., Wu, N., Zhu, H., Teng, L., Zhang, W. (2017). SVM-DT-based adaptive and collaborative intrusion detection. *IEEE/CAA Journal of Automatica Sinica*, 5(1):

- 108-118. <https://doi.org/10.1109/JAS.2017.7510730>
- [29] Pawar, P.S., Hashmi, S.A. (2015). Security enhanced adaptive acknowledgment intrusion detection system. *International Journal of Computer Applications*, 975: 8887. <https://doi.org/10.5120/ijca2015907055>
- [30] Wafa'S, A.S., Naoum, R.S. (2009). Adaptive framework for network intrusion detection by using genetic-based machine learning algorithm. *IJCSNS*, 9(4): 55.
- [31] Lee, W., Stolfo, S.J., Mok, K.W. (2000). Adaptive intrusion detection: A data mining approach. *Artificial Intelligence Review*, 14: 533-567. <https://doi.org/10.1023/A:1006624031083>
- [32] Kumar, C.K., Ramachandran, N. (2022). A comprehensive review on intrusion detection and prevention schemes for network coding enabled mobile small cells. *Ingénierie des Systèmes d'Information*, 27(1): 29-39. <https://doi.org/10.18280/isi.270104>