REDUCING NETWORK TRAFFIC WITH A NET FLOW ANALYZER IN WIRELESS SENSOR NETWORKS

Vasanthakumar S[1], Ranjith N[2],

ABSTRACT:

Network traffic reduction in a Net Flow Analyzer tool involves the process of optimizing data collection and analysis to focus on relevant information while minimizing unnecessary network data. Implement filters to capture only specific types of traffic, such as filtering by source or destination IP addresses, port numbers, or protocols. This reduces the volume of data collected. To aggregate the data and summarize information, such as by creating flow records that consolidate multiple packets into a single record, aggregation reduces the volume of data without losing essential details. Time-based analysis is the analysis of network traffic based on specific time intervals or patterns. Despite constantly watching every component of the network configuration, this procedure aids in the recognition of abnormalities. upcoming sites for aggregation are forthcoming able to execute increasingly intricate collection of information procedures due to significant advancements in very low-power processor performance, which will lessen the vulnerability of WSNs.

Keywords: Network Traffic, Filter Data Aggregate, Secure Data Aggregate Techniques, Threat Model, Detecting Collusion Sub Aggregate Alerts, Robust Data Aggregation, Iterative Filtering Algorithm.

I. INTRODUCTION

Applications for Associations of electronic sensors are also referred to as providing a growing number, including military surveillance, detection of forest fires, and monitoring of wild habitats. During the array of hops involving your foundation facility, sensors are located autonomously. Serving for instance, the hub after they are placed in the area of interest. A sensor node's computational power and energy storage are typically severely limited. Allowing its control centre to receive measurements from every sensing component. potentially through an additional intermediary the simplest means of gathering perceived information gathered by an internet connection requires centres to handle newly acquired information earlier than a foundation facility does. But in terms of transmission overhead, this approach is unaffordable (or energy expended). By integrating partial findings at intermediate nodes during message routing or computing aggregates in-network, big WSNs can drastically cut communication and, consequently, energy consumption. Several data collection solutions for WSNs employ the technique of building a spanning tree with its roots at the base station, followed by in-network aggregation around this limb. Both major research efforts have relied on overall aggregation methods. These methods take global network conditions into consideration. It should be noted that such aggregates can be easily extended to include predicates and conditions. For example, aggregation may consider only those sensor readings that exceed a threshold value, such as measurements greater than one hundred units.Moreover, count and sum can be used to calculate the average. Every empirical instant plus average, variance, can additionally be calculated using an extension of the sum technique. Strategies of gathering involving trees have proved resistant towards switching; therefore, transmission failures, which frequently occur in networked wireless sensors (WSNs), result in interaction losses. The research community has suggested using several journey directing algorithms to carry sub-aggregates in order to solve this issue. This method offers a flexible approach towards replication-insensitive mergers like Min and Max. Nevertheless, varied track direction causes the detectors to estimate values twice for duplicate-sensitive aggregates like Sum and Count. A number of researchers have recently proposed ingenious solutions to address the multi-path approach-related double-counting issue. The summary diffusion framework is a scalable and resilient aggregation approach that was developed when trying to compute duplicate-sensitive clumps, like Sum and determine. With this method, a node in the aggregation may have more than one parent due to the ring topology. Structure, and a duplicate-insensitive bitmap known as a synopsis represents each detected value or sub-aggregate.

Department of Computer Science,
Karpagam Academy of Higher Education, Coimbatore, Tamilnadu, India
vasanth.chml@gmail.com

Department of Computer Science,
KSG College of Arts and Science, Coimbatore, Tamilnadu, India
ranjithksg@gmail.com

* Corresponding Author

II. LITERATURE OF REVIEW

The hidden irregular behavior within the system can reduce its overall detectability. Network traffic data from each router within the campus network is collected, processed, aggregated, and stored in a centralized NetFlow log file. According to Yang [1], authentic and up-to-date network flow patterns are essential for developing effective strategies to detect network intrusions. Damascus, R. [2]. In order to gather information about road traffic as well as available parking spots in a smart city, we suggest an effective mechanism for controlling congestion that is based around sensor-based system construction. Hillman, A [3]. The velocity method of control reduces the delay from end to increase network life over long simulation periods. The hybrid K-means and greedy best first search algorithms are used for cluster nodes at first. Srivastava [4]. This is made up of small sensors that communicate with the Internet of Things to record and track physical conditions. The sensor nodes are self-contained and build an ad-hoc collaboration pattern using one another. Hayseed K. [5]. The software-defined networking paradigm emerged as an interesting way to implement alternative routing control strategies, opening up new possibilities for the shipment of distributed Internet for Everything apps. In order to enhance sustainability within WSN, interaction is the avenue scalp that guarantees payments lacking sleeping spots upon wiring are able on-condition equipped point through locate towards shipment. Theodora T., Justus J.J. [6–7]. With the radio medium's presenting nature, OR techniques address two major issues in bound networks of wireless sensors. chaotic the node adaptability while attributes correspond quality. For greater distribution alignment within its computer system, OR has the ability to decrease delays. Chithaluru P. [8]. Our gadget has multiple branches in the region, comprising paths, antiques, and intersections, and all of them interact and exchange commute flow and automotive data. It possesses a few overcrowded courses in terms of distance and time. Prasad, J. P. [9]. Receive more modifications during the match philosophy power distribution technique is proposed to increase the overall network lifetime. It has been demonstrated that the modified GTEB outperforms the current algorithms in the following areas: duration ratio, scheduling overhead, figure of hops, energy exhaustion, belief of active motes, figure of decaying molecules, and productivity HV. [Chair]. The two main factors affecting the battery life of low-load wireless sensors are data flow and point-state switching. Although the node's switching energy cannot be disregarded, a decreased node burden results in low energy use during data transmission. Zing. [11]. Since data compression reduces the amount of data that must first be transferred by the observing root to the network's collapse node [12]. In [15], the researchers showed how hierarchical structures combined with machine learning can further address VANET challenges in a way conceptually similar to the original paper's hierarchical dissemination model and emphasized the importance of intelligent cluster head selection, QoS-aware routing, and machine learning techniques for efficient communication and data dissemination in VANETs.

III. METHODOLOGY

The model operates in multiple stages using batches of consecutive sensor readings. In the first stage, it provides an initial estimation of variation, taking into account noise parameters associated with the sensors. The calculations used in this stage help estimate each sensor's bias and variance. A novel approach is introduced to utilize sensor data for estimating noise variance and bias.The distortion amplitude and bias of a sensor can be interpreted as measures of how much the sensor readings deviate from the true signal value. In fact, even in the case of non-stochastic (systematic) errors, the estimated measures of dispersion and skewness serve as reliable indicators of instrument variability and bias.

A. Breach of data privacy:

Every exploited location information that operates for their compatible intelligence accumulator might release measurement information (indeed, subordinate stones) that its opponent receives through its siblings.

B. Collusion of the sub-aggregate:

If node C is hacked, it can collide with the sub-aggregate, which is calculated using messaging that other cells send to parents.

C. Collusion local value:

That hacked node B might spoof with another measurement examining for affecting every single score. Because the prevailing rate for a genuine station carries whatever (i.e., is unaffected by its scope of activity), this hacked path might say it actually feels something. In that instance, a cooperative regional exploit cannot be identified.

D. Consolidation Portion:

Every router sends several recognition packets and carries out an amalgamation element that is a creative narrative spreading procedure. It ought to have been noted that the actual bonded description BX estimated for the intersection X might range from the actual synthesized storyline BX throughout the collaborating sub accumulation

exploit.

IV. MATERIALS AND METHODS

1. Secure Data Aggregation Techniques.
2. Threat Model.
3. Detecting Collusion Sub Aggregate Alerts.
4. Robust Data Aggregation.

A. Strategies for Safe Statistics Gathering:

This verification technique founded on trees proved created so that an underpinning unit could ascertain which specialize total-count along with sum-was the result of collusion. This concept cannot be extended for synopsis verification because the synopsis computation is duplicate-insensitive. A verification technique was created to calculate the sum and count in its spreading of narrative method. The method we employ is quite similar. Employing that exception takes a different approach to try and prevent your costs for interaction even more. Furthermore, we offer thorough theoretical analysis to determine the optimal balance between communication overhead and security. New strategies to use the "secure contracted gathering" exhibit have been recently established. However, those techniques intended for use with wireless sensor networks on the other hand, if there is just one malicious node, this technique is secure. When there are a few compromised nodes, the computing count and sum during the SDAP attestation phase can be costly. Based on a sampling technique, its assault-resistant Determine with defined amalgamation procedure computations is currently suggested. This approach is able to yield an estimate concerning the intended total in spite of adversarial interference. We had earlier provided the merger mechanism regarding the summarization propagation paradigm being durable for attacks. The reader may have some queries because the topic covered in this article is more specific than the one covered in our prior work. We emphasize that while our prior work tackles a broader issue, it has a significant delay and lacks a lightweight verification algorithm.

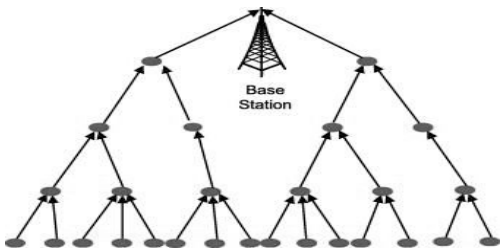

Figure 1 : Proposed overall architecture

B. Threat Model :

There are no security features in the summary diffusion framework by itself. Consequently, it's at risk many types of engages from unapproved hacked hubs. It is possible to add common authentication and encryption protocols to the aggregation framework in order with the aim safeguard unapproved ports against listening in interfering with conversations amid reputable locations. Therefore, since adversary can gain cryptography keys from the compromised nodes, we do not see the necessity to take into account assaults that originate from unauthorized nodes. Cryptography procedures are unable to stop attacks that are conducted by infected circuits.

C. Private information violation:

fortunately, an infected additionally functions alongside the compatible statistics marketplace. it is possible for the adversary to obtain sensor readings and sub aggregates that are received by the nodes that are their offspring. Privacy-preserving algorithms were proposed by several researchers.

D. Collusion of the local value:

In order to affect the aggregate worth, vulnerable point may these instances were identified by collusion's particular scanner.

1. Assuming an prevailing rate significant sincere site remains unlimited (i.e., unaffected within their app region), then infected server may argue its existence feels anything. Under such a scenario, conspiring regional exploit cannot be identified.
2. Once a legitimate cluster has boundaries while defined as penetrated root distorts stay inside this obligated, then nobody has nothing for identifying this kind of assault.
3. Someone upfront site's state has boundaries, while stolen point contradicts its neighborhood value irrespective a bounds constraint.

E. Collusion of sub-aggregate:

It is difficult to defend against this attack since a hacked node has the ability to alliances, actual sub-aggregate, was intended for calculation using a notification that get delivered through parent nodes.

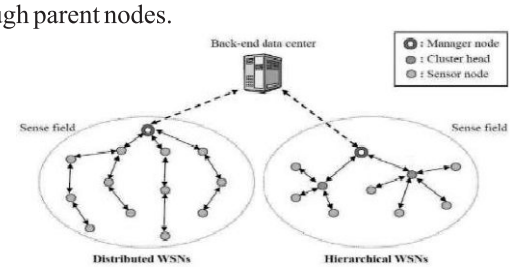

Figure 2 : Thread Model

F. Detecting Collusion Sub Aggregate Alerts:

Since a sinking entity establishes their assemble prediction upon lowest-order got r, meaning 0 with the last integrated outline, a targeted devices might have created the unique merged overview before being able alter a quantity

regarding r. These could be accomplished with ease with showing your caregivers the resulting bound breakdown along with entering several bits in spaces j, that $r \leq j \leq k$. Remember offering what every impaired point needs just adjust the computation for r where the recede generates involves setting several greater-order fragments regarding 1. details hacked point lacks the capacity to be aware of r's genuine worth.

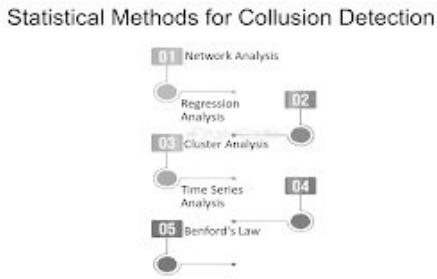

Figure 3: Methods for Collusion Detection

From one warped description, consider r′ stand alongside the lowest-order gently which had been set too 0, and r for called lowest-order gently which gets assigned on 0, with an appropriate explanation. aspects sink's cumulative prediction is thereby 2r′ larger over the precise figure. Through employing this previous method, it's clearly evident this vulnerable node can add the important degree extra variance to the final sink measurement. the summation diffusing mechanism may use by essentially one device for launching the offensive having excellent results due to a method utilizes several paths navigation, which improves chances a conspiring interpretation gets to transmitter.

G. Robust Data Aggregation:

Innovative method for calculating the noise variance and bias for sensors based on their measurements. One way to interprets distortion along with volatility during circuits vibration is a measure of how much the sensor readings deviate from the actual signal value. As a matter of fact, for non-stochastic errors, the distance measures that we arrived at as estimates of the bias and variances of sensors also make sense. This allowed us to Calculate an variety surrounding the accurate representation about something that was assessed, thus constitutes just one variable through an probability purpose.

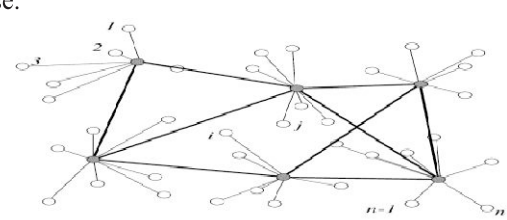

Figure 4: Data Aggregation Layout

The attacker takes advantage of the IF algorithms' flaw, which results from a false presumption about initial reliability of the sensors. Using a robust data aggregation technique as the foundational principle for these algorithms is our contribution to addressing these problems. Furthermore, the distance between a sensor's reading and such an initial reputation can be used to calculate the initial weights for each sensor node.

IV. NETWORK SIMULATOR

With a wide range of Internet protocols, including those for satellite, wireless, and terrestrial networks, NS is a public domain emulator. The most widely utilized simulator for research papers that are presented at prestigious conferences such as SigComm is NS. Its enormous user base and a small team of ISI developers work tirelessly to maintain and upgrade NS. A discrete event simulator designed for networking research is called NS. An optional network animator is included. Ns and its partner, Nam, make up an extremely potent combination of teaching resources for networking topics.

A. Starting Nam:

Two ways of starting Nam are from within a The Toshiba virtual reality characters. In a modelling, yourself wish to understand, I could begin it 'Nam's' directive in hand," where " represents called Actual record title that NS generated. A screenshot of a Nam window with an explanation of its most crucial features may be found below.

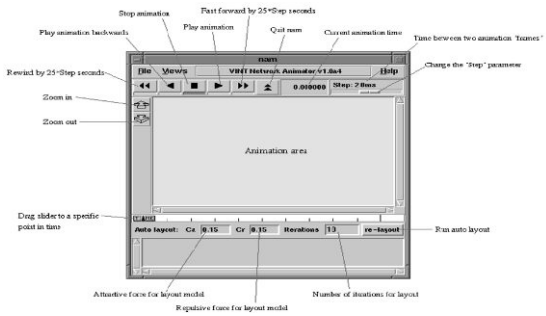

Figure 5: Network Animating Window

B. Iterative filtering Algorithm:

Iterative filtering algorithms Typically, this type of aggregation takes the form of roughly that is biased. indicators with results.

Step 1: Network construction.

Step 2: Roofing in this region.

Step3: Determine the neighbour's separation while generate suitable navigation.

Step4: Deciding the Flexible Longitude Between Every framework.

Step5: Decompose the nodes into a particular area (location).
Step 6: Accumulating vertices based on their region.
Step7: Demand modifications on every point.
Step 8: Transferring information is required an achieve safe multi -hop relaying.
Step 9: Manufacturing slots-based data aggregation flow and filtering for data.
Step10: Transmitted data and slots to the receiver.
Step 11: Estimating the area-based stream of information removing impurities.
Step 12: Handle Conduct Upkeep.
Step13: Reverse on knowledge give alongside their path.
Step 14: Calculating energy.
Step 15: Evidence shippers are re-adjusted after an email catching exhausted takes place collective.

C. Testing

An amount during independently pathways across block of its main conduit scripts constitutes a content measure of its cyclomatic complexity. For example, the complexity would be 1 That exists merely as a single strategy of parsing its genome, while an original script contained neither choice rewards, such as checks, nor until repetitions. Mathematically, a directed graph with an edge connecting two fundamentally programmed blocks is used to define the cyclomatic complexity of a structured program. If the source code has a single path and a cyclomatic complexity of 1, it does not contain a control flow statement first and second fundamental building blocks, since control can flow from one to the other.

Thus, the definition of cyclomatic complexity M would be: M is equal to E compared to N is defined as + 2A, when M is the tropical intricacy. Details graph's guard count (E), node density (N), while linked parts count (P) are expressed as follows.

V. RESULTS AND DISCUSSION

Understanding the items that are modelled in software and the relationships that link them is the first stage in black box testing. Making a graph of the imp. objects and their relationships are the first step in testing. Next, a set of tests covering the graph is designed to exercise each object and relationship and identify any problems.

This allows us to use the following techniques for behavioural testing:

1. Modelling transaction flow
2. Modelling in finite states
3. Information is the modelling of flow.

To make sure that all objects and their relationships are executed, we now need to focus on node coverage and link coverage**.**

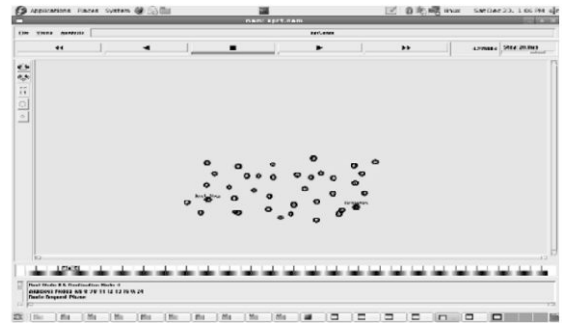

Figure 6: Rooting Construction

Figure 6 is Routing simulators, network analysers, and auditing tools constitute some of the tools that are often utilized in establishing networks for path inspection and optimization. Diverse network monitoring tools, ranging from Nemesis or deliberately Solar Winds Limited additionally to packet tracers for emulating network layouts, are instances of analytical tools. Wire shark is used for packet analysis.

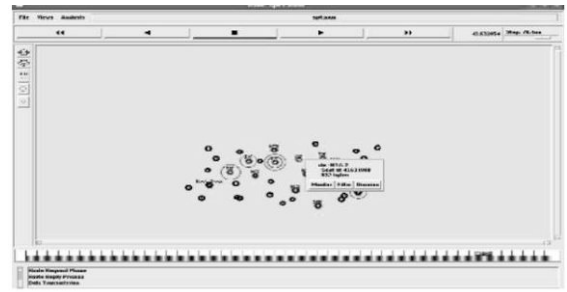

Figure 7: monitor node creation packet files

Figure 7 Choose a strategy for locating and capturing node creation files and packets. Protocol analysis, packet sniffing, or integration with current network infrastructure may be required for this. Place sensor nodes in the network in strategic locations to gather pertinent data. Take into account variables like conversation range, node density, and coverage.

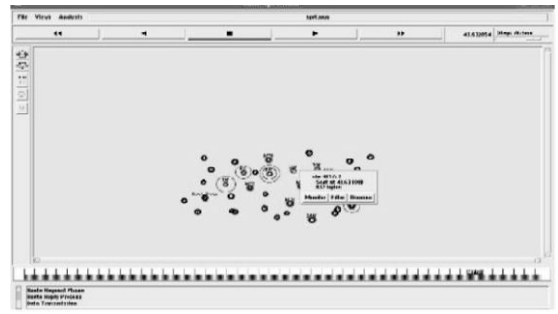

Figure 8: Data Communication

Figure 8 is to get information about network flow, and use sensors. Information about device-to-device communication, such as source and destination IP addresses, ports, protocol, and communication duration, is often included in network flow data. Make use of a WSN's wireless communication capabilities to send gathered data to the machine along with a crucial collection point.

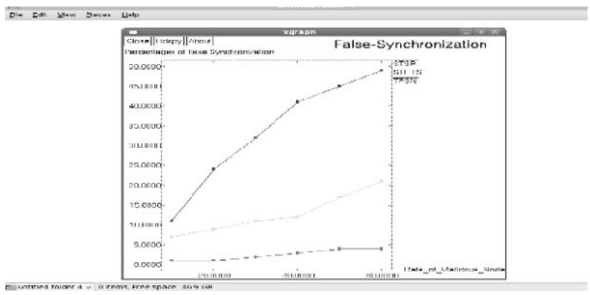

Figure 9: False Synchronization

Figure 9 shows It is possible to analyse network traffic patterns, spot bottlenecks, and improve communication routes by using a graph-based method. Regarding analysing networks, graph algorithms such as prominence initiatives and methods of clustering, including shortest path methods, may be used.

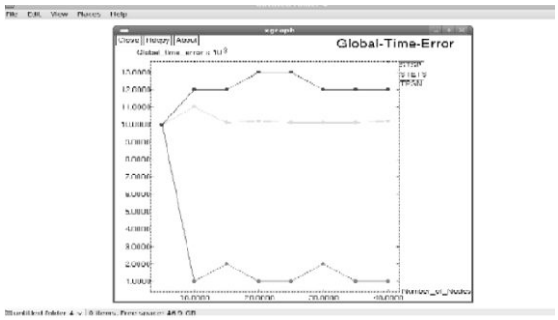

Figure 10: Global Time Error

Figure 10 Make sure that timestamp configurations across all the gadgets and infrastructure pertinent are set to coincide if you are experiencing problems with timekeeping, including time-consuming inaccuracies in a regional setting. Defects in traffic analytics and NetFlow analysis are just two of the many issues that can arise from inaccurate time settings.

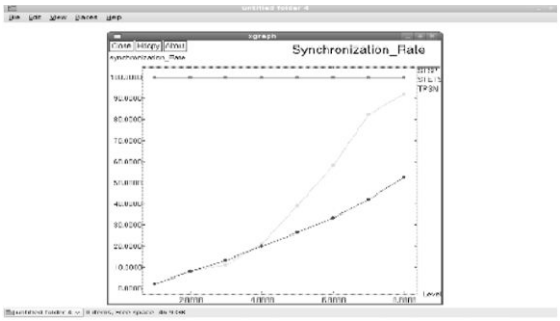

Figure 11: Synchronization Rate

Figure 11 is The Gephi protocol is an open-source program for graph visualization and analysis in broad networks. For viewing the interactions during node locations, machinery, especially corners, or affiliations, customers may incorporate circuit metrics and apply numerous arrangements of algorithmic reasoning.

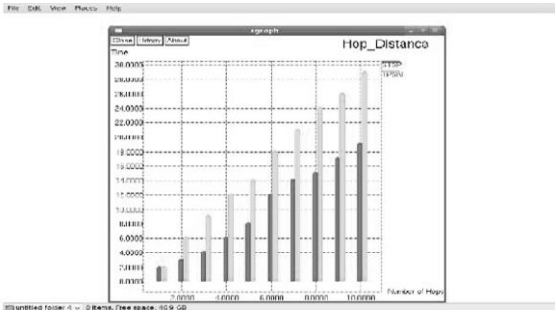

Figure 12: Hop Distance

Figure 12 is Although this kind of aggregate alert can be used with many other distributed systems, it is especially risky when used against wireless sensor networks (WSNs) for two reasons. Secondly, node-compromising attempts are quite likely to target sensors.

VI. CONCLUSION

Traffic reduction is the process of optimizing data collection and analysis to minimize extraneous network data and concentrate on pertinent information. This decrease improves the tool's functionality and increases its ability to offer network filter data insights. They stated that there is a considerable danger of false positive schemes, making it impossible to revoke hacked nodes that have been found. One of the primary characteristics of this strategy is the employment of several monitoring nodes for carrying out aggregation functions and supplying a MAC quantity corresponding to the collected results as a component for the desktop computer in the figure computed by the cluster-wide integrator

REFERENCE

[1] Javari, Daniel, "Fuzzy logic-based Dados attacks and network traffic anomaly detection methods: classification, overview, and future perspectives." Information Sciences, 2023.

[2] Majid, Mammon, "Applications of wireless sensor networks and internet of things frameworks in the industry revolution 4.0: A systematic literature review." Sensors 22.6, 2022.

[3] Chaitra, H. V., "Delay optimization and energy balancing algorithm for improving network lifetime in fixed wireless sensor networks." Physical

Communication 58, 2023.

[4] Yao, Chengpeng, "Traffic Anomaly Detection in Wireless Sensor Networks Based on Principal Component Analysis and Deep Convolution Neural Network." IEEE Access 10, 2022.

[5] Awed, Mohammed, "Examining the suitability of NetFlow features in detecting IoT network intrusions." Sensors 22.16, 2022.

[6] Gupta, Amit, and Adesh Kumar. "Study on the wireless sensor network routing for low-power FPGA hardware in field applications." Computers and Electronics in Agriculture, 2023.

[7] Gawky, Mohammed Zaid, "Research Article: An Effective Wireless Sensor Network Routing Protocol Based on Particle Swarm Optimization Algorithm.", 2022.

[8] Diaz, Alvaro, and Pablo Sanchez. "Modelling various threats regarding safeguarding via smart sensor platforms.", 2016.

[9] Ponnusamy, Vasaki, "IoT Wireless Intrusion Detection and Network Traffic Analysis." Computer Systems Science & Engineering 40.3, 2022.

[10] Yadav, Saneh Lata, "Traffic and energy-aware optimization for congestion control in next-generation wireless sensor networks." Journal of Sensors, 2021.

[11] Krishnaveni A, Arvinth K A, “Smart Animal Monitoring System using YOLO and OpenCV, “Proceedings of the International Conference on Electronics and Renewable Systems (ICEARS 2025), 11-13, February, 2025.

[12] Mythili, S., and G. Karunakaran. "A Hybrid Approach for Content Filtering and Blacklisting in OSN." In 2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), pp. 1141-1147. IEEE, 2024.

[13] M. M. Karthikeyan, “Building Resilient API Security Through a Five-Dimensional Framework for Data Breach Prevention in Modern Digital Ecosystems”, Partners Universal Multidisciplinary Research Journal (PUMRJ), vol. 2, no. 4, 2025.

[14] R Sivakumar, VB Kirubanand, V Ganesan, M Sivaraman, A Kumarachelvan, G Ulaganathan, "Building Smarter Systems with Advanced Computational Techniques," International Conference on Inventive Computation Technologies (ICICT), 772-777, Publisher IEEE, 2025.

[15] Krishnakumar, K. G., and E.J.Thomson Fredrik, “QoS enabled data dissemination in hierarchical VANET using machine learning approach”, International Journal of Signal and Imaging Systems Engineering, Vol 10, Issue 5, 2017.