AN INTRUSION DETECTION APPROACH USING GRAPH-BASED DEFENCE MECHANISM

N. Mohanasundaram[1] *P. Sherubha*[2] *S. P. Sasi Rekha*[3]

ABSTRACT

Wireless sensor networks (WSNs) have created a huge interest among investigators in the last few years involving numerous challenges. This huge interest is significantly connected to diverse applications modelled on large scale networks comprising devices' competency to carry out evaluation towards sensed data and process data by transmitting them to remote locations. WSNs have been provided with security that plays a crucial role among network systems and deployed generally in unreachable terrain and for communication towards wireless domain. The outcomes provide security methods that are injected into extremely vulnerable sensor networks that are vigorous to deal with attacks from malicious nodes. WSN comprises nodes with constraint resources and traditional security measures that are accessible to conventional networks that are not used here. Therefore, the necessity of using those systems relies on sensor nodes boundary resource and has competency to deal with these attacks. Modelling the network with more appropriate defence mechanism has an ability to identify unknown/malicious attacks and find solution to thwart them using graph-based defence mechanism (GDM). Henceforth, intrusion detection maintains an extremely effectual research field for investigators. Thus, familiarity with this research field may resourcefully beneficial for investigators. Simulations have been done in MATLAB environment. The anticipated method shows better trade off compared to prevailing approaches.

Keywords: *Wireless Sensor Networks, malicious node, graph theory, defence mechanism, constraint resources*

I INTRODUCTION

Wireless Sensor Networks (WSNs) are measured as distributed wireless networks which consist of a huge amount of lower cost sensory nodes that are deployed in monitoring field [1]. Sensor nodes (SNs) with certain resource constraints such as computational capacity, memory and energy are associated with untrusted and unreliable networks [2]. In recent times, WSN with guaranteed needs has been elaborately utilized in military and commercial domains. Moreover, conventional security approaches are inconsistent to fulfil security needs owing to WSN nodes-based limitations [3]. Henceforth, meeting the security needs may turn out to be a very challenging factor.

As an essential approach to fulfil network security, intrusion detection technology has been constantly utilized as baseline for WSN-based defence mechanism [4]. The ultimate target of Intrusion Detection System (IDS) is to identify intruders that attempt to disrupt those networks or to examine WSN security and recognize vulnerability to fulfil precise network performance. Investigators attempt to optimize node proportion for detecting intrusion during node deployment. Similarly, in [5] the author attempts to utilize offline detection for detecting intrusion and utilize intrusion detection to protect WSN from routing attacks. Certain anomaly-based detection tries to identify anomalies by describing ordinary characteristics. Intrusion detection may help administrators to discriminate 'normal nodes' from 'abnormal nodes' owing to abnormal characteristics utilizing classification procedures. In [6], the author attemptsto examine PDR anomalies and receives signal strength indicator to identify a jammer. In [7] the author targets to identify network graph discrepancy against routing attacks.

[1] Professor, Department of Computer Science and Engineering, Karpagam Academy of Higher Education, Coimbatore. Tamilnadu. India.

[2] Research Scholar, Department of Computer Science and Engineering, Karpagam Academy of Higher Education, Coimbatore. Tamilnadu. India.

[3] Research Scholar, Department of Computer Science and Engineering, Karpagam Academy of Higher Education, Coimbatore. Tamilnadu. India.

Due to similarity among IDS and biological immune system, intrusion detection approach based on Artificial Immune System (AIS) has been investigated in computer networks and associated fields [8]. AIS isdivided into four diverse families: Clone selection, Artificial Immune network, and negative selection and Danger theory. AIS-sourced Holland's classifier is anticipated for network intrusion detection. AIS with social learning are utilized to quicker convergence speed and to eliminate anti-bodies that les within local optimal for optimization purpose. Danger theory is utilized to diminish false alarm in four diverse layer model of IDS. In danger theory, dendritic cells areutilized to develop four-layer models for network IDS. Negative selection algorithm (NSA) isanticipated initially by Forrest et al. [9] in 1990s and resourcefully simulated with immune tolerance procedure of human immune system for recognizing non-self and self approach. It is utilized to resolve anomaly and fault detection crisis. AIS is effectual to resolve potential crisis in WSN applications owing to WSN which is extremely similar to biological body in system features and system composition with features such as adaptability, self-learning, self-organization, memory pattern, robustness and so on[10].Moreover, there are certain constraints when NSA is utilized in WSN-based IDS owing to nodes with restricted resources, for instance, battery power, restricted memory space and computational ability.

So as to deal with the crisismentioned above, this work anticipates IDS of WSN as WSN-based defence mechanism,which is dependent on diverse intrusion detection approach with enhanced Graph-based defence mechanism that adopts certain optimizing factors for diminishing time-based cost, memory consumption, computing cost and so on. The ultimate contribution of GDM over WSN is to provide an immune system specifically utilizing this graph-theory-based modelling to automate WSN-based intrusion detection. The baseline objectives are provided here:

1) Modelling a pre-processing setting for WSN for effectual communication with other devices using graph theory concept. This works as a system model for reducing memory and cost computation.

2) Modelling a heuristic attack strategy for fulfilling the basic needs of sensor network for offering self-adaptive model for speeding up radius and for handling real time response for diminishing attack identification time and cost.

Modelling a multi-level design for reducing computational power consumption where optimization factors are deployed in base station, collecting normal data for deployment with ordinary nodes specifically for detection purposes.

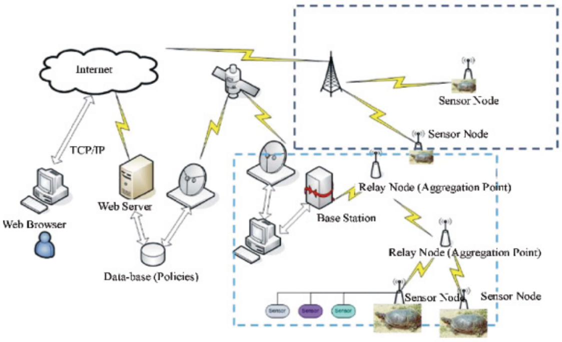

Fig 1: Generic WSN architecture

The remainder of the work is organized thus: Section II is a background study, Section III anticipated graph model with intrusion detection, Section IV specifically for numerical results and discussionsand Section V provides conclusion with advice for future extension.

II RELATED WORKS

Certain clusters in WSN display foremost and effectual characteristics while implementing diverse functions termed as Mobile-based Wireless Sensor Networks (MWSN). In recent times, WSN-based mobility has turned out to be an interesting field of research. Transportability is basically measured to hold diverse crises that need to be resolved such as coverage, connectivity and energy consumption [11]. To provide certain factors, WSN deployment is not are considered to be entirely stationary. Subsequently, modern and present studies show a broader perspective in sympathetic and positive light[12]. Moreover, mobility authenticates sensor nodes to pursue a travelling approach for instance, transportation, chemical clouds and packages. The significant ability of preserving protective range from fire boundary along with recent information and updation regarding fire fighters focuses on boundary in certain time constraint.

While discussing WSN, it is necessary to pursue a major distinction between mobile WSN and static WSN from diverse factors such as dynamic network topology; localization, power consumption and network sink [13]. Various comparisons depict numerous benefits of mobile and static WSN in comparison to replication attacks. The comparison provides performance characteristics of MWSN in contrastto WSN for localization purpose, power consumption, dynamic network topology and network sink. In certain systems it is placed inside static region, and dead sensor nodes may man oeuvre to merge weaker communication and observation paths. Using this statistical WSN, this approach is not attainable as information from cut off, dead or detached nodes may primarily disappear [14]. Simultaneously, when positioning sinks with immobility, sensor nodes are placed nearer to base station that may die rapidly. Therefore, it may transfer huge data for communication and information to nodes that are nearer, that is, auxiliary apart. This crisis can be solved and network lifetime may be upgraded with the use of mobile base stations [15]. The essential benefits of mobility-based WSN are in contrastto WSN to offer superior coverage, higher channel capacity and enhanced tracking and enhanced energy efficiency.

III PROPOSED METHODOLOGY

This section describes the system model of network design. Consider $G = (V, E)$ is measured to the original network model, in which $V = \{v_1, v_2, \dots, v_n\}$ is set of nodes, $E = \{e_1, e_2, e_3, \dots, e_n\}$ is cluster of links connected with network. Attack encountered in network is related to connection and disconnection of links which is specified as $Attack = \{+e_1, -e_2, \dots, +e_{2r-1}, -e_{2T}\}$ is solution of compromising link series, where $+ \, and \, -$ are addition and removal of links respectively towards the network. Then, network is depicted as trails in Eq. (1):

$$\begin{cases} V' = V \\ E' = E + attack \end{cases} \tag{1}$$

Therefore, it is essential to determine such attack to modify network connection and acquire adversarial network. For this purpose, G', community detection approach is drastically worse, that is, the quality of detection outcomes decreases.

a. Attack strategy

Here, link-based attacks may change certain relationship in node connectivity. Here, link-based attacks are rewired to preserve targeted nodes' degree that is, including link towards target nodes when deleting it from connectivity. Those kinds of rewiring are considered being of a significantly lesser cost, that is, only two nodes are essential for rewiring attack, where some nodes may change degree if addition and removal of links are encountered. Re-modifying depicts hiding the real link from the original target node. Here, most of the links maintain the same originality. Some random attacks are considered in this work. Initially random attack is considered here. Nodes are chosen randomly from G to form target node V. Then, randomly choose node $V_i \in V_t$ for all iterations and execute following operation: erasing existent link when adding non-existent one from target node. Neighbourhood node is provided by $N_i = \{v_j | < v_i, v_j > \in E\}$, while non-neighbourhood is chosen by $\overline{N_\iota}$. Total amount of remodified attacks, provided as $'T'$, is another factor that specifies attack cost. Random strategy is provided below:

Network with certain structural model offers certain characteristics like higher density of intra-community links and lesser density of inter-connecting links. Adding and eliminating links within the community may result in weakening the network structure. With this heuristic attack strategy merged with results via certain community detection approach is termed as community-based attack detection. In case of target nodes, v_i inter-community nodes are set as $I_1 = C_i \cap N_i$, where C_i is community node set. Alike of inter connected non neighbour set $S_2 = \overline{C_\iota} \cap N_i$, where $C_i = V - C_i$,

It is determined that network connectivity may provide degree-based power distribution, constant with traditional $80/20$ rule. Generally, lower number of nodes may hold greater number of connections while the rest are determined as real-world networks. Certain modifications encountered in hub nodes may have higher degree of impact towards network structure. For example, people with more friends in social media may have higher circle. Hence, heuristic attack strategy may be termed as Degree-based attack that targets attack nodes with huge network degree.

The significant difference among these two attack types are chosen based on network degree as targeted nodes randomly. Therefore, choosing target nodes has network degree while others may remain idle.

Algorithm 1

```
Input: $T, K, G$
Output: $G'$
        $V_t \rightarrow random\ attack\ (G, K)$
Attack count = 0;
For attack count < T do
$v_i \rightarrow$ random count $(v_i)$
$N_i, \overline{N_\iota} \rightarrow$ set $(G, V_i)$
If $N_i \neq \emptyset$ and $\overline{N_\iota} \neq \emptyset$ then
        $V_{delete} \rightarrow random\ (N_i)$
$v_{add} \rightarrow$ random $(\overline{N_\iota)}$
Eliminate $< v_i, v_{del} >$ from E and add $< v_i, v_{add} >$ to E;
Update G;
Attack count = count +1;
End
End
Construct adversarial network for re-connecting attacks.
```

Algorithm 2

```
$C = \{C_1, C_2, .., C_h\} \rightarrow$ detection $(G)$;
For $length\ (K - node) < K$ do
            $v \rightarrow degree\ (V)$;
        $V = V - C_i, where\ v \in C_i$;
  If $K > h\ then$
  Revise V each turn;
  Repeat from biggest degree;
  End
  End
```

b. Continuous Event Measures

Volume of intrusion detection environment and frequency in monitoring environment are evaluated incessantly. Here, Gaussian model exemplifies continuous data event, where probability density function is provided by:

$$P\ (v|\sigma, \mu, k) = \sum_{i=1}^{k} w_i .\ N(v|\mu_i, \sigma_i) \qquad (2)$$

Where $'v'$ specifies normal event value, $'k'$Gaussian components, $'m'$ and $'s'$mean and standard deviations, andw_i component weight. Gaussian model is computed with Expectation Maximization. The number of components is done with Bayesian criterion for selection.

For every incoming event $\emptyset_j$, EM algorithm is computed as probability of v_j to deviate Gaussian model. Probability value of v_j will be smaller or larger than maximum or minimal data instances. Computation details are depicted in three phases.

In the initial phase, Gaussian element nearer to $v_j$in Gaussian model is depicted as k^* component. Distance measure is utilized, where distance between v_j and k^* Gaussian element is evaluated by as:

$$h_k^*(v_j) = \frac{|v_j - \mu_k *|}{\sigma_k^*} \qquad (3)$$

In the second phase, distance normalization by standard data instances belongs to k^* Gaussian component is evaluated as:

$$y_m = \frac{h_k^*(v_j) - \mu_m}{\sigma_m} \tag{4}$$

In the third phase, probability to deviate k^{th} Gaussian element is evaluated as:

$$p(v_j \geq \max v \,||v_j \leq \min v) = e^{-e^{-ym}} \tag{5}$$

Last step is anomaly score computation of events to translate probability by:

$$\propto (\emptyset_j) = \min(\frac{-\ln(1-p)}{\tau}, 1) \tag{6}$$

Where, τ is the highest anomaly score for normalization. However, anticipated approach is extendedly inherent to assist event with multi-variate values.

c. Intrusion detection

Intrusion detection works towards security improvement from H2H communication among nodes' agent in sensor infrastructure [15]. So as to recognize malicious error propagation and insertion, sensor ID works in two phases in sensor infrastructure offered by larger transmission range. Subsequently, it identifies compromised nodes in sensor network.

Unauthorized node identification works in network initialization phase. It is initialization of neighbourhood vehicles that is accountable for authentication of unknown nodes or new nodes in network. Consider, new node that joins in network with prevailing authenticated nodes, Neighbouring nodes in communication range is joined in network for authentication. Consider node that is out of coverage range, network node initialization under direct transmission range is provided as below:

$$X_1 \rightarrow B, C: m_1 \tag{7}$$

$$B, C \rightarrow X_1: m_2 \tag{8}$$

$$X_1 \rightarrow B, C: m_3 \tag{9}$$

$$B, C: m_4 \tag{10}$$

$$B, C: m_5 \tag{11}$$

Where encryption parameters are provided for data security, distribution function is used for determining confidence interval. Authorized network initialization and neighbouring nodes have wo diverse constraints that have to be fulfilled.

$$m_4 == m_1 \tag{12}$$

$$m_5 == m_2 \tag{13}$$

Once network initialization is completed successfully, it further helps in initializing newer nodes by joining network. Suppose there are some authenticated nodes and some other nodes want to join the network,one node is out of transmission range and the other node moves towards transmission range. However, this process is done for subsequent ranges.

Node detection works during H2H dissemination by considering infrastructure nodes' moving capability owing to availability or regularity in traffic environments. Detection is done with confidence computation for next hop with distribution function.

IV Numerical results and discussions

Experiments were carried out in MATLAB to analyze the performance of anticipated graph-based intrusion detection approach to evaluate its ability with other prevailing models. In this investigation, time-series-based data is produced in intervals. Consider randomly distributed sensor nodes that collect uni-variate data with respect to room temperature and time instants. Fig 2 demonstrates sensor placements configuration and normalized sensors with maximum readings.

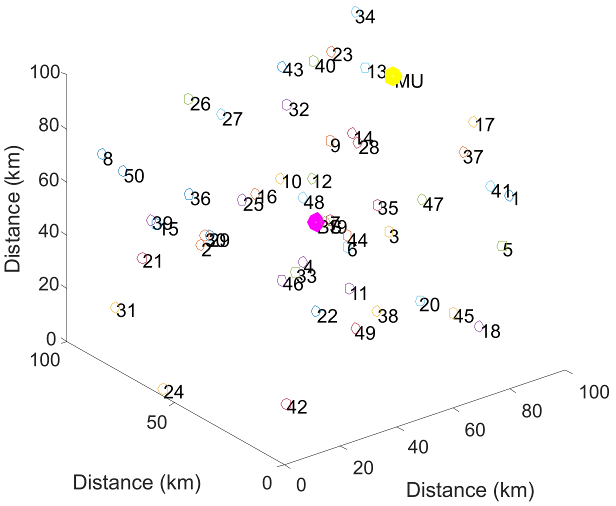

Fig 2: Node Establishment

Intelligent intrusion detection systems can be constructed only if the effectual dataset isaccessible. The dataset imitates the real time data, and can help in testing and training an intrusion detection system. NSL KDD set is an updated version of conventional predecessor KDD'99 data set as in Table II. In this work, NSL-KDD data set will be examined and utilized to investigate the efficiency of diverse classification algorithms in anomalies detectedin network traffic patterns. This dataset comprises quality of data.

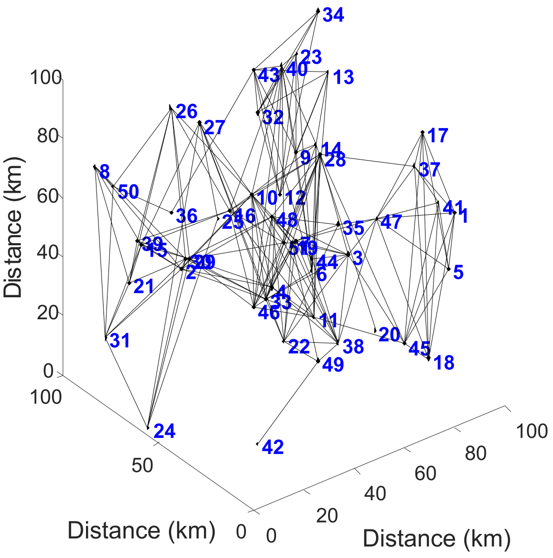

Fig 3: Graph Connectivity

NSL-KDD data set is examined and classified in four different clusters illustrating four common types of attacks as given below. An in-depth analytical investigation is done on training and testing data set. The performance and execution speed of diverse clustering procedures are examined. Here 20% of training and 80% of data set are used for testing as in Table I. Here, NSL-KDD data set is used to project vulnerable protocols that are frequently utilized by intruders to launch attacks. Table II shows the parameters related to NSL-KDD data set.

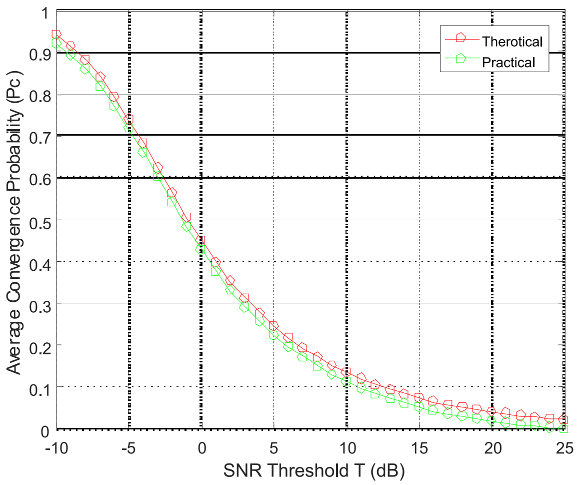

Fig 4: Coverage Probability

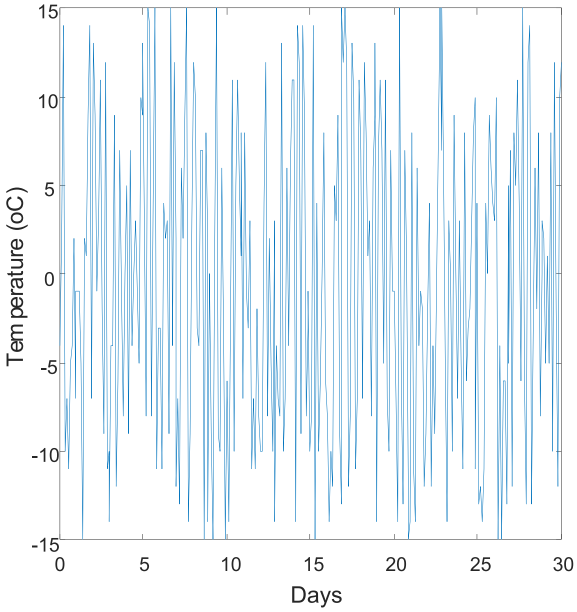

Fig 5: Node Temperature

Confusion matriX	Normal	Attack
Actual	-	-
Attack	21	736
Normal	227	16

Table I: Confusion Matrix

KDD training	125973	67343 / 53.46%	45927/36.46%	995/0.79%	52/0.04%	11656/9.25%
KDD testing	22544	9711/43.07%	7458/33.08%	2754/12.22%	200/0.89%	2421/10.74%

Table II: NSL-KD Training And Testing Factors

Att_No.	Att_Name	Explanation	Sample Data
1	Duration	Time duration length of connection	0
2	Protocol_type	Protocol utilized in establishing connection	TCP
3	Service	Utilization of d estination network service	FTP_data
4	Flag	Connection status – Normal	SF
5	Src_bytes	Sum of data bytes from S to D	491
6	Dst_bytes	Sum of data bytes from D to S	0
7	Land	If both IP addresses /port numbers are similar, variable considers 1/0	0
8	Wrong_fragment	Total amount of incorrect fragments	0
9	Urgent	Sum of urgent packets .It is packets with urgent bit activated	0

Table III: General Dataset Attributes

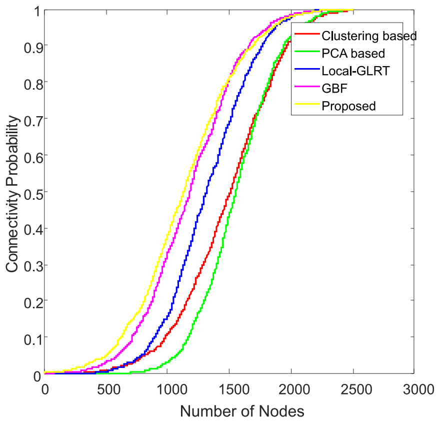

Fig 6: Connection Establishment

To comparethe computational complexity of anticipated approach withother techniques, we compute essential CPU time with average runs, when experiments are evaluated in MATLAB on Intel Core with 4 GB RAM. Average CPU times are needed for diverse intrusion detection approaches. The running time of process the graph based approaches is computed; however its detection rate is drastically superior as shown in Fig 4, Fig 5 and Fig 6 to this approach.

V. CONCLUSION

WSNs have been provided with security that plays a crucial role among network systems and deployed generally in unreachable terrain and for communication towards wireless domain. The outcomes provide security methods that are injected into extremely vulnerable sensor networks that are vigorous to deal with attacks from malicious nodes. WSN consists of nodes with constraint resources and henceforth traditional security measures that are accessible to conventional networks are not used here. Therefore, the necessity of using those systems that relies towards boundary of sensor nodes based resource potential and eligible to deal with these attacks. Modelling the network with more appropriate defence mechanism may help to identify unknown/malicious attacks and find solution to thwart them using graph-based defence mechanism (GDM). Henceforth, intrusion detection may be an interesting research field for investigators.

References

[1] M. Ahmed, A. N. Mahmood, and J. Hu, “A survey of network anomalydetection techniques,” Journal of Network and Computer Applications,vol. 60, pp. 19–31, 2016.

[2] Z. Liao, Y. Yu, and B. Chen, “Anomaly detection in GPS data based onvisual analytics,” in VAST’10, 2010, pp. 51–58.

[3] R. A. Leite, T. Gschwandtner, S. Miksch, S. Kriglstein,M. Pohl, E. Gstrein,and J. Kuntner, “Eva:

Visual analytics to identify fraudulent events," IEEETransactions on Visualization and Computer Graphics, vol. 24, no. 1, pp.330–339, 2018.

[4] F. Beck, M. Burch, S. Diehl, and D. Weiskopf, "A taxonomy and surveyof dynamic graph visualization," in Computer Graphics Forum, vol. 36,no. 1, 2017, pp. 133–159.

[5] J. Zhao, N. Cao, Z. Wen, Y. Song, Y. R. Lin, and C. Collins, "#FluxFlow:Visual analysis of anomalous information spreading on social media,"IEEE Transactions on Visualization and Computer Graphics, vol. 20,no. 12, pp. 1773–1782, 2014.

[6] W. Jiang, C. Gu, and J. Wu, "A Dynamically Reconfigurable Wireless Sensor Network Testbed for Multiple Routing Protocols," Wireless Communications and Mobile Computing, vol. 2017, Article ID 1594270, 10 pages, 2017.

[7] A. Xenakis, F. Foukalas, and G. Stamoulis, "Cross-layer energy a ware topology control through Simulated Annealing for WSNs," Computers and Electrical Engineering, vol. 56, pp. 576–590, 2016.

[8] F. Wu, C. R¨udiger, and M. R. Yuce, "Real-time performance of a self powered environmental iot sensor network system," Sensors, vol. 17, no. 2, p. 282, 2017.

[9] T. N. Le, A. Pegatoquet, O. Berder, and O. Sentieys, "A Power Manager with Balanced Quality of Service for Energy-Harvesting Wireless Sensor Nodes," in International Workshop on Energy Neutral Sensing Systems(ENSSys). Memphis, United States: ACM, Nov. 2014, pp. 19–24.

[10] F. Ait Aoudia, M. Gautier, and O. Berder, "GRAPMAN: Gradual Power Manager for Consistent Throughput of Energy Harvesting Wireless Sensor Nodes," in IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications, Hong Kong, China, Aug. 2015,

[11] P.-D. Gleonec, J. Ardouin, M. Gautier, and O. Berder, "Multi-source Energy Harvesting for IoT nodes," in IEEE Online Conference on Green Communications (Online GreenComm), Nov. 2016.

[12] Xiaoyang Liu 1,2 and Chao Liu, "Wireless Sensor Network Dynamic Mathematics Modeling and Node Localization", Hindawi Wireless Communications and Mobile Computing Volume 2018, Article ID 1082398, 8 pages https://doi.org/10.1155/2018/1082398

[13] MarwaKazdoghliLagha, Fayc¸al Ait Aoudia, "Feature Selection Framework for Multi-source Energy Harvesting Wireless Sensor Networks", IEEE 2011.

[14] Krzysztof Lasota, Piotr Bazydło, Adam Kozakiewicz, "Mobile Platform for Threat Monitoring in Wireless Sensor Networks", European Union, 2016.

[15] Unai Burgos 1,2, UgaitzAmozarrain, "Routing in MobileWireless Sensor Networks: A Leader-Based Approach", Sensors, 26 May 2017; Accepted: 3 July 2017; Published: 7 July 2017.