VOLUME 1, 2018
DOI: 10.31058/j.data.2018.22003

A Method for Identifying a Wormhole Attack Using Regression in Wireless Sensor Networks

Seyyedjalaleddin Dastgheib[1*]

[1]Department of Computer Engineering, Shiraz University, Shiraz, Iran

Email addresses

dastqeib@gmail.com (S.J. Dastgheib)
*Correspondence: dastqeib@gmail.com

Received: 17 July 2018**; Accepted:** 15 August 2017**; Published:** 29 August 2018

Abstract:

Due to the specific features of wireless sensor networks (WSNs), such as a large number of nodes and limited energy, we need to use unique methods for this type of network in dealing with attacks. There are numerous attacks on WSNs that cause loss of network security. One of these attacks is setting the wrong path using a wormhole. A wormhole attack is triggered by two adversary nodes in such a way that the packets are broadcasted by an adversary to another adversary without decoding or separating each packet. Adversaries are directly linked to each other by a vast transmission range. In this thesis, we try to detect and neutralize this attack using regression. This paper presents a comprehensive approach that, in all conditions, without the need for additional equipment and with the highest accuracy and low energy consumption, can be detected and ultimately countered by the attack of the wormhole.

Keywords:

Wireless Sensor Network, Security, Wormhole Attacks, Regression

1. Introduction

Because of the unique features of wireless sensor networks (WSNs) such as the large number of nodes, the limited power supply of nodes, etc., the issue of dealing with security attacks in these types of networks has become one of the fundamental challenges, because despite the similarity of the type of attacks with other wireless networks, there is a need for unique solutions based on the specific features of WSNs [2].

There are several types of attacks that compromise the security of the WSN. One of these attacks is setting the wrong path using a wormhole. In a wormhole attack, two adversary nodes create a transfer channel between themselves. The endpoints of this channel are called start and end points. The adversary node sends its data through the channel to the adversary node at the endpoint [3]. Then the endpoint broadcasts the received packets in its covered area. So it creates the impression that the nodes are neighbors even though they are far apart from each other. In addition to routing, the wormhole attack can easily demolish the localization by altering the number of

neighboring nodes, channeling the anchor nodes, resizing the communication range among nodes, and so on [4].

The location information is broadcast by anchor nodes on the network. Anchor information is transmitted improperly in case of falling into a wormhole. Sensor nodes are waiting to receive location information from nodes in wormhole attack areas within a given time frame. But because of the wormhole, they do not get this information, and they suspect that there is a wormhole. In this case, the nodes analyze the records of packet rates received from their neighbors. In this paper, a/the regression technique is used to investigate the occurrence of the wormhole. Regression means to return, to predict and express variations of a variable based on the information of one or more other variables. At first, regression information is completed at the training phase, and then, if a node is suspected of the wormhole, the packet entry log information is given to regression. The output indicates the occurrence or absence of wormhole. We expect this method to be very carefully designed to detect and eliminate the wormhole with minimal energy consumption.

The remainder of this article is organized as follows: Section 2 examines the previous methods for identifying wormhole attacks. In Section 3, the proposed wormhole attack detection is introduced. In Section 4, the proposed method is compared with several suitable and robust schemes in the field of wormhole attack discovery and neutralization. And Section 5 concludes the paper.

2. Related Works

Robust position estimation (Rope) is an independent domain-based layout based on the two-level network architecture that provides passively distributed localization [6]. The sensors determine their own location by referring to the wires that are received from locators, with specific coordinates and directions. Each locator is equipped with a directional antenna, so it covers a different sectional area with different transitions. In each sector, the coordinator sends out its coordinates, and the domain of the boundaries of the sector being transmitted refers to the commonly known axis.

An analysis of connection graphs to detect wormhole attacks is used in [7]. In addition, the false negative is considered to be much more dangerous than the false positives (some legal connections will be suspected of having a glue). When a false positive connection is removed, a valid connection will be lost, but security will not be compromised. A false negative, on the other hand, keeps the network unsafe. The authors of the paper have attempted to eliminate suspicious connections when the number of neighbors of a node suddenly increases significantly, in order to deal with wormhole attacks. The structure of the connection graph is one of the main challenges of this form; in addition, the detection and removal of connections that do not impose a great cost on the network.

Secure Localization with Mobile Beacons (SLMBs) is a securely centralized sensor locating scheme using a beacon (anchor) node [8]. It uses an animated beacon node to move between a calculated path in the network to collect spatial information associated with anonymous sensor nodes. The information is then sent to the base station, where the coordinates of the sensor nodes are calculated. If the beacon node from a given position receives more than one node or receives location information from a node more than once, it suspects the existence of a wormhole.

In [9] authors have provided a safe way to locate distance vector hop (DV-Hop) using method [9]. In the DV-Hop methods, the distance between the two nodes is determined by the middle hops (steps) between the two nodes, and with the wormhole, this number of steps is under affected. This paper analyzes the critical effects of wormhole attacks on location based on the DV-Hop. To address this security issue, the schema based on the label is presented to define and defend against wormhole attacks for the DV-Hop localization process.

In [10] the exploratory view of wormhole attacks has been done comprehensively. Though a different number of countermeasures are defined to deal with this wormhole, almost all of them suffer from disadvantages that affect the large-scale WSN. A wormhole geographic distribution diagnostic algorithm (WGDD) is manufactured to use the DV-Hop technique as a method for the testing wormhole.

In [11] a mechanism for detecting and preventing wormhole attacks is provided. In this way, no specific hardware is required. All that is done is to calculate the round trip time (RTT) of each path to calculate the RTT threshold. Regarding the simulation results of various parameters such as the average latency to the end, the packet delivery section and the average permissive power, it has been proven that the proposed mechanism works better than the wormhole affected by the Ad hoc On-demand Multipath Distance Vector (AOMDV) multiple-range interpolation vector routing protocol.

The topological differences created by wormholes, the network analysis, and retrieval based on the multidimensional scale (MDS) using RTTs to detect wormhole bonds are presented in [12]. This method can detect multiple wormhole bindings connected by short and long paths. In this paper, the wormhole diagnostic method is presented using a domain-based topology comparison that exploits the local neighborhoods.

There are also review articles on attacks on WSNs that can be mentioned [13].

3. The Proposed Method

In the proposed method, the distribution of nodes is considered as a two-dimensional, flattened environment. Distribution of nodes is considered in a randomized manner. The number of anchor nodes and the number of nodes in the WSN are adjusted so that localization is done without a problem. The number of adversary nodes in the wormhole can be more than two, resulting in more wormhole paths than one. The hypothesis of this article is the existence of a wormhole and two adversary nodes, but if the attacker's nodes increase, they will be able to adapt to this condition.

The proposed method presented in this paper is a fully distributed method, meaning that all operations involve the detection and counteraction of the wormhole by the nodes itself, and the sink does not play a role in this. On the other hand, additional equipment is not used on the network, and the entire node has unique physical and hardware properties. In other words, the network is homogeneous. We examine the proposed method in two phases of identification or discovery and neutralization phases.

3.1. Identification Phase

In the identification phase, as mentioned, sensor nodes are responsible for analyzing and identifying wormholes. After the wormhole develops, actually changes occur in network traffic, and some nodes receive a greater volume of information, while in some nodes the amount of received information is reduced.

Each attacker (adversary) node in the wormhole affects the area of the network and attempts to deceive nodes in the area. This is a circular area with an attacker communication radius as its radius and attacker position as the center point. The radius of an attacker cannot be infinite and is usually slightly larger than the sensor node's transmission range. If the range of sensor nodes is Rc, the RA is considered as the attacker communication radius (area) in this research. The attacker communication radius is a radial in which an attacker re-launches another attacker's message. In addition to the node's radius of communication and the attacker's radius, there is also the communication radius of an anchor node or locator that is represented by the RL. We have in total:

$$Rc < RL < RA \quad (1)$$

There are the nodes that participate in the diagnosis of wormholes. These nodes will see the greatest change in case of wormhole occurrences. The nodes within the scope of attackers in the wormhole are in this category. The nodes on the margin of the wormhole cover also see changes. Before the wormhole happens, packets of nodes that are now covered with wormholes are received to marginalized nodes. After the wormhole is created, the nodes in the wormhole's area of coverage send packets to the wormhole, resulting in fewer packets being sent out of the wormhole's coverage area. This reduces the number of packets received by nodes on the margin of the wormhole.

Changes in the number of packets received by the node help to discover the wormhole. In this paper, the number of packets received is controlled by each node. This control is done in such a way that the number of packets received is calculated in specific time periods. Because of low memory, only the current time and previous time information are maintained. In other words, only packet information received from existing angles is stored in two periods of time, and the older information is cleared by recording new ones. By analyzing the packets received at the current time in a node, the probability of occurrence of wormholes can be realized. As we know, in the event of a wormhole, packets sent to a node are significantly different from the time before it. After the wormhole is created, the nodes in the wormhole region receive a large number of packets from the attacker position, which is due to the broadcast packets of the nodes located in the area of endpoint adversary.

At an angle where the wormhole is located, it may have received no information or the information that it receives is negligible. In the event of a wormhole, the information received by the node from this angle, while the node in the area covered by the wormhole, is much greater than before. In order to ensure greater accuracy, adjacent angles are also examined to avoid computational errors. The wormhole also affects the side angles, especially if the node is covered in the adjacent area.

In the adjacent nodes covered by the wormhole, the information is sent to the node before the onset of the attack from a specific angle, which will significantly reduce the incidence of wormholes. These nodes also use the same mechanism to explore the wormhole. The difference in packet reception rates from an angle in the node at the current time with this rate is a prime criterion for detecting the wormhole by the node at a previous time.

Each node of the sensor, during the main operation of the network, maintains the closed arrival rate of its neighbors from the neighboring angles. We assume that the sensor node receives the information from 5 nodes. The node in the normal environment, which does not significantly change packet arrival rates, assumes that the wormhole does not exist and follows the normal network execution process. This is logical because the effect of a wormhole attack is more on the packet stream and the packet traffic pattern changes in the network.

The difference in packet arrival rate is defined by the angles of the feature vector. For example, (5, 4, 3, 5, and 6) is the arrival packet rate in current time from neighboring nodes of node i, and (5, 4, 4, and 5, and 7) is the packet arrival rate in previous time from node i neighbors. In this case, the feature vector will be in this example (0, 0, 1, 0 and 1)=2. There is not much difference in this vector. Therefore, if the difference in arrival packet rates from the neighbors of the current node is 0, 0, 1, 0, and 1, then the wormhole has not occurred.

Here is a regression to identify a wormhole. But before that, learning has to be done. In the regression method, the feature vector, where is the difference in packet arrival rate, is given as an input, and then determines the output of the wormhole's existence or absence. In the learning step, creating a distinctive feature vector adds a row to the training table, and the result of the existence or absence of a wormhole is placed in a separate class. For example, see Table 1.

Table 1. *Regression learning.*

Differences in packet arrival rate	Class of existence or absence of wormhole
a:(0,1,2,0,1)= 4	N
b:(1,3,14,3,19)=40	Y
c:(2,0,2,3,1)=8	N

If the learning table has not grown sufficiently, the training process should continue. Therefore, a row is added to the table, in which one cell is the difference in angle of arrival. A specific process is used to determine the presence or absence of a wormhole and to insert it in the other cell of the table, as described below.

In the correlation analysis, the primary goal is to measure the linear correlation between the two variables, but basically, we do not seek such measurements in regression analysis. We try to estimate or predict the average value of a variable. The regression equation is as follows:

$$\hat{y} = \alpha + \beta x \tag{2}$$

In this formula, α is the width of the origin and basic coefficient, and β is the gradient (slope) of the regression line and the coefficient of the angle (regression coefficient). For each variable unit in variable x, the variable y changes to β.

$$\beta = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2} \tag{3}$$

$$\alpha = \frac{\sum y - \beta \sum x}{n} \tag{4}$$

A sensor node receives information from neighboring nodes in its neighborhood. In the event of a wormhole, the data exchange in the WSN is impaired. To prevent the wormhole's impact on the network, wormholes should be identified and neutralized as soon as possible. Sensor nodes normally hold packet arrival rates in previous time

intervals and average packet arrival rate. Due to the memory limitation, we consider *k* as the number of time intervals for maintaining the packet arrival rates.

If the packet arrival rate at the current time in a node has a significant difference with the packet arrival rate at the k previous intervals from the same angles, the node is suspected of the occurring wormhole. The packet arrival rate from a different angle can be considered as an elementary array, where n is the number of angles that the packet arrival to the node. To obtain a similarity between two arrival rates, the Jaccard similarity function is used to determine the similarity of two arrays. We use Jaccard's similarity [14] as a similarity function. Jaccard's similarity function is shown for calculating the similarity of two arrival rates at t and t-1 time with the symbol corr (t, t + 1), and is defined as follows:

bi(t) = {rate1, rate2, ..., raten}

$$Corr_{t,t-1} = \frac{b_i(t).b_i(t-1)}{\|b_i(t)\|_2^2 + \|b_i(t-1)\|_2^2 - b_i(t).b_i(t-1)} \tag{5}$$

$$Where\ \|b_i(t)\|_2^2 = |rate_1|^2 + |rate_2|^2 + ... + |rate_n|^2$$

For simplicity, k = 2 is considered, and the following equation is used to suspect the wormhole.

$$\begin{aligned} &\text{If } corr_{t,t-1} < \text{threshold} \\ &\text{and } corr_{t,t-2} < \text{threshold} \\ &\text{and } corr_{t,avg} < \text{threshold} \Rightarrow \text{suspects to WORMHOLE} \end{aligned} \tag{6}$$

The node, if suspected of being a wormhole, wants other neighboring nodes to examine the wormhole in accordance with the above relationship and return the result. If more than half of the nodes positively to wormhole occurrences, the node is sure of the wormhole, and it performs its neutralization mechanism.

3.2. Neutralize Wormhole Phase

The wormhole discovery process described in the previous section will greatly facilitate the neutralization of the wormhole. The node that recognizes the wormhole occurs, according to the angle of incoming analysis and the received signal, in addition to the detection of the wormhole, can calculate the spatial information around it. This location information includes the angle of the wormhole position relative to that node and the distance between the node and the wormhole.

The process of neutralizing the wormhole happens with this local information available. The node that detects the wormhole disperses this information outside the affected area. The nodes that observe this information prevent the transmission of their data to the wormhole. The nodes around the wormhole also misplace and do not select the other attacker to send the packet. As a result, the attacking nodes go to isolation and actually get out of work.

4. Results and Discussion

This section should present your findings objectively, explaining them largely, concisely, precisely in the text. You should show how your results contribute to the

scientific knowledge in academic community clearly and logically. Results and Discussion may be divided by subheadings, concluding equations, figures, tables, etc.

In this section, we will evaluate the proposed method and compare it with the WGDD [10] and MDS [12] methods. Among the previous designs for the detection of wormholes, these two methods are the most authoritative ones. As a result, proving the superiority of the proposed method on them can demonstrate the effectiveness of the proposed method.

To begin the simulation, the primary parameters of the network must first be identified. The initial parameters presented in Table 2 include the number of nodes; the distribution of the nodes, the primary energy of the node, the dimensions of the node distribution environment, the sink coordinates, etc.

Table 2. *Simulation primary parameters.*

Parameter	Value
Number of Sensor Nodes	200
Network Length	100
Network Width	100
Sink Coordinates	(50,50)
Initial Energy	0.5 Joule
Packet Length	32 bit
Nodes Deployment	Randomly

To evaluate the proposed method, we consider that the sensor nodes are randomly distributed in a flat and two-dimensional 100x100 network.

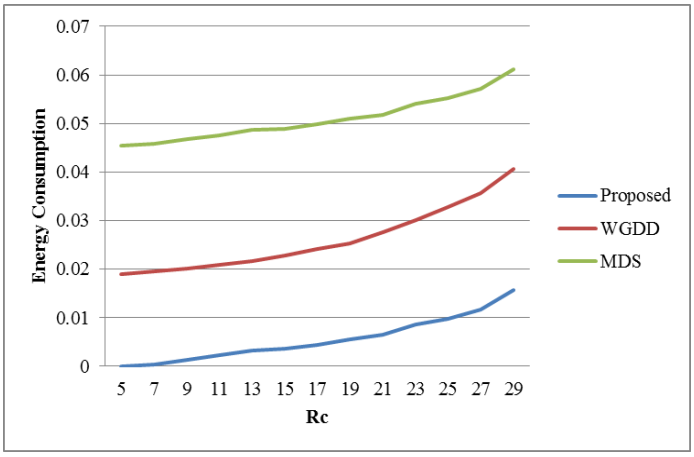

Figure 1. *Energy Consumption per Rc in a network with 50 nodes.*

In Figure 1, energy consumption is used to detect wormholes in WGDD, MDS and proposed methods in the event that the number of nodes is 50. The significant difference between the energy consumption of the proposed method with the WGDD and MDS methods is shown in these graphs. We do this simulation for different communication radii (RC). The communication radius of Figure 1, which is located on the horizontal axis, starts at 5 meters and extends up to 29 meters at intervals of 2 meters. All results for this communication radius are considered at each simulation time, and the indicated graphs are obtained.

According to the results, the energy efficiency of the proposed method compared to the WGDD and MDS methods in terms of changing the range of transmission has a remarkable advantage, and the proposed method always provides better energy consumption.

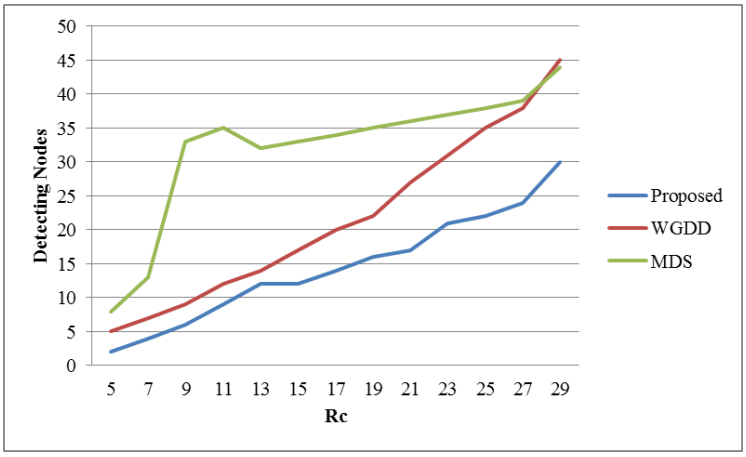

Figure 2. *The nodes involved in the process of detecting for different Rcs with 50 nodes.*

The next issue we are dealing with is the number of nodes involved in the process of detecting and dealing with a wormhole. The smaller the number of nodes, the better the performance of the proposed method, because it reduces energy consumption and imposes less overhead on the network. In Figure 2, when the number of nodes is 50, the number of nodes involved in the proposed method, WGDD and MDS, have been investigated. As it can be seen, the number of nodes involved in the proposed method is far less than the number of nodes involved in the other two methods.

We examined the effect of changing the communication radii of sensor nodes. Regarding Figure 2, in the different transmission domains, the proposed scheme has a lower number of nodes than WGDD and MDS. This issue itself is evidence of the superiority of the proposed method on WGDD and MDS.

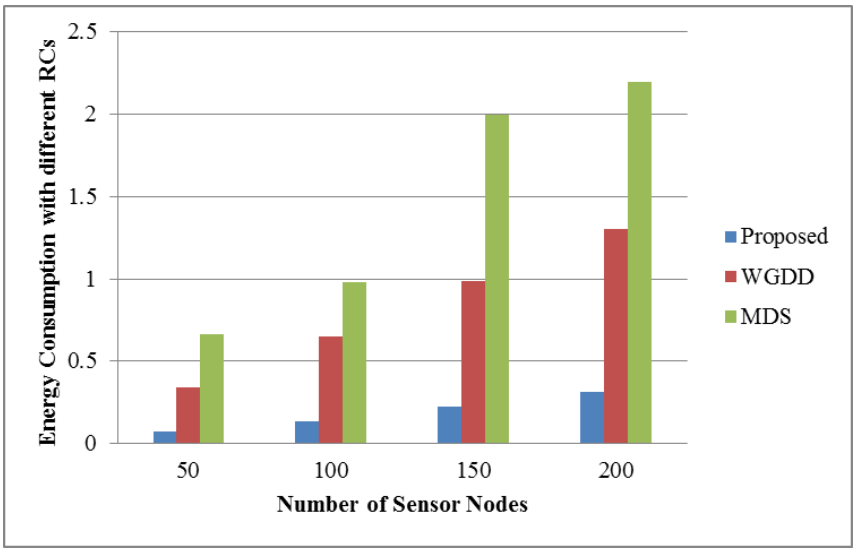

Figure 3. *Total energy consumption per various Rs.*

Figure 3 shows a comparison of energy consumption in the proposed method with WGDD and MDS methods, in which the number of nodes varies from 50 to 200 with an increase of 50. According to Figure 3, energy consumption in the proposed method has been significantly reduced by WGDD and MDS methods.

In addition to a general review of energy consumption in different conditions, including a number of sensor nodes and different RCs, the number of nodes involved in the process of detecting and counteracting wormholes was also analyzed. To

determine the superiority of the proposed method, we consider WGDD and MDS methods in terms of the number of nodes involved to examine them together. To do this, we calculate the total number of nodes for different RCs in each node of the network with a given node number. What you see in Figure 4 is the comparison of the total number of nodes involved in the process of identifying and dealing with wormholes for a different number of network nodes. Based on this figure, it is clear that the proposed method has fewer involvement nodes in all modes than the WGDD and MDS methods.

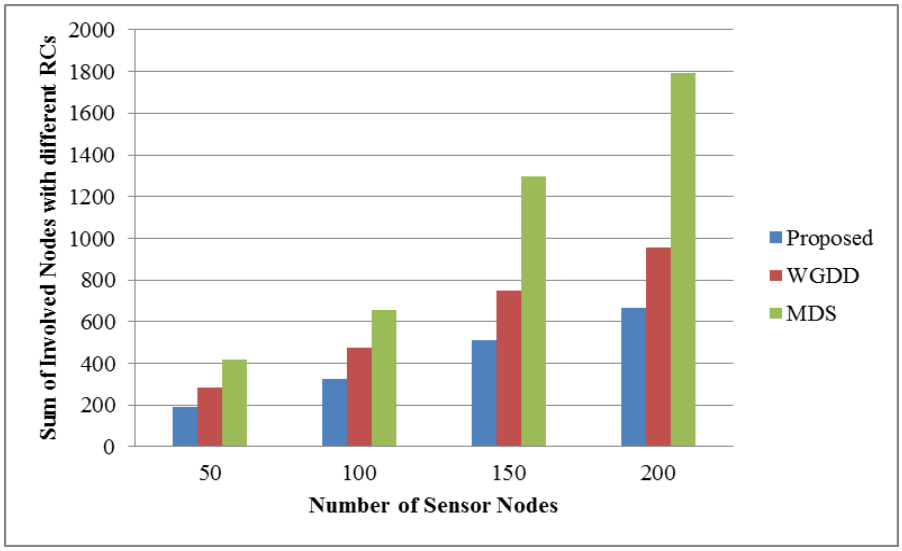

Figure 4. *Total involved nodes per various Rc and sensor nodes.*

5. Conclusion

The attack we investigated in this study is a wormhole attack on the localization process in a WSN in which the wormhole causes a mistake in localization of the nodes by spatial redirection of spatial information in the network. A wormhole is fully distributed, detected and neutralized by sensor nodes itself, using regression designed for nodes. The idea of the wormhole's discovery process is inspired by the analysis of data arrival from different angles because the wormhole re-circuits the information part of the network to another part and affects the flow of information. The benefits of this design can be distributed, using regression, reducing the nodes involved in the wormhole exploration process and reducing overhead. To demonstrate the effectiveness of the proposed scheme, this scheme is simulated and the results are presented in the previous section.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this article.

References

[1] Marques, B.F.L.G. Application-Driven Wireless Sensor Networks. Doctoral dissertation, Universidade do Porto, 2017.

[2] Charalampidou, M.; Pavlidis, G.; Mouroutsos, S.G. A novel modular wireless sensor networks approach for security applications. *International Journal of Security and Networks*, 2017, 12(1), 40-50, DOI: 10.1504/ijsn.2017.10001807.

[3] Sabri, Y.; ElKamoun, N. GRPW-MuS-s: A Secure Enhanced Trust Aware Routing against Wormhole Attacks in Wireless Sensor Networks. *Communications on Applied Electronics (CAE)*, 2016, 6(5), 1-9, DOI: 10.5120/cae2016652472.

[4] Chen, H.; Lou, W.; Wang, Z.; Wu, J.; Wang, Z.; Xia, A. Securing DV-Hop localization against wormhole attacks in wireless sensor networks. *Pervasive and Mobile Computing*, 2015, 16, 22-35, DOI: 10.1016/j.pmcj.2014.01.007.

[5] Liu, L.; Luo, G.; Qin, K.; Zhang, X. An algorithm based on logistic regression with data fusion in wireless sensor networks. *Journal on Wireless Communications and Networking*, 2017, 2017(1), 10, DOI: 10.1186/s13638-016-0793-z.

[6] Lazos, L.; Poovendran, R.; Čapkun, S. ROPE: robust position estimation in wireless sensor networks. In Proceedings of the 4th IEEE international symposium on information processing in sensor networks, 2005, 1-43.

[7] Ban, X.; Sarkar, R.; Gao, J. Local connectivity tests to identify wormholes in wireless networks. In Proceedings of the Twelfth ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2011, 1-13.

[8] Zhang, T.; He, J.; Yu, H. Secure localization in wireless sensor networks with mobile beacons. *International Journal of Distributed Sensor Networks*, 2012, 8(10), 197-189, DOI: 10.1155/2012/732381.

[9] Dabi, P.S.; Tunwal, K.; Khandelwal, R.; Acharya, D. A Survey of Detection of Wormhole Attacks in Wireless Sensor Network. *International Journal of Enhanced Research in Management & Computer Applications*, 2014, 3(6), 52-55.

[10] Ughade, S.; Kapoor, R.K.; Pandey, A. An Overview on Wormhole Attack in Wireless Sensor Network: Challenges, Impacts, and Detection Approach. *International Journal of Recent Development in Engineering and Technology*, 2014, 2(4), 102-109.

[11] Rajeswari, S.R.; Seenivasagam, V. Comparative study on various authentication protocols in wireless sensor networks. *The Scientific World Journal*, 2016, 2016(1), 1-6, DOI: 10.1155/2016/6854303.

[12] Amish, P.; Vaghela, V.B. Detection and prevention of wormhole attack in wireless sensor network using AOMDV protocol. *Procedia computer science*, 2016, 79(1), 700-707, DOI: 10.1016/j.procs.2016.03.092 .

[13] Mukherjee, S.; Chattopadhyay, M.; Chattopadhyay, S.; Kar, P. Wormhole Detection Based on Ordinal MDS Using RTT in Wireless Sensor Network. *Journal of Computer Networks and Communications*, 2016, 2016(1), 1-15, DOI: /10.1155/2016/3405264.