Australian Journal of Basic and Applied Sciences, 9(20) June 2015, Pages: 90-95

ISSN:1991-8178

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

Intrusion Detection for Wsn Using Automata Theory

[1]N.Vadivelan and [2]S. Anbu

[1]*Research Scholar, Department of Computer Science and Engineering, St.Peter's University, Avadi, Chennai, Tamilnadu, India.*
[2]*Professor, Department of Computer Science and Engineering, St.Peter's College of Engineering and Technology, Avadi, Chennai, Tamilnadu, India.*

ARTICLE INFO

Article history:
Received 16 April 2015
Accepted 12 June 2015
Available online 1 July 2015

Keywords:
Security in WSN, Intrusion Detection System, Finite Automata Expression.

ABSTRACT

Providing security for Wireless Sensor Network-[WSN] is more essential nowadays. Data transmission among two nodes in WSN happens through multi-hop nodes. Since, it is necessary to analyze the intermediate nodes is much important to verify the nodes are trustable nodes or not. In this paper, an Expression verification method is applied for detecting the intruder by analyzing the packet transmitted in a specific route. The personal information, the behavior of the nodes and the data are represented in FAE-[Finite Automata based Expression] format. The nodes' states are matched with the states of the automata for various expressions especially timer based timeout message. The memory utilization by the FAE reduces the resource requirements, less memory and it can provide more accuracy in state changing and expression verification. The simulation is carried out in Network Simulator and the performance is evaluated by repeating the experiment for various numbers of nodes deployed in the simulation environment

To Cite This Article: N. Vadivelan and S. Anbu., Intrusion Detection for WSN using Automata Theory. ***Aust. J. Basic & Appl. Sci.,*** *9(20): 90-95, 2015*

INTRODUCTION

A Wireless sensor network (WSN) is a temporary network composed of a collection of wireless sensor nodes, it requires no existing infrastructure or centralized administration. Wireless sensor network construction and deployment is cheap, easy and location independent. WSN-based communications are used in applications from healthcare to military. They are currently used for monitoring human activity, natural disasters and environmental phenomena. WSN is intensively used for emergency military situations and surveillance monitoring. Since wireless sensor nodes are mobile and dynamic, it is difficult to send data both safely and with the efficient use of energy. The battery levels of the nodes must be considered with regard to routing and the frequent mobility of the nodes. Since each node in the network has a limited wireless transmission range, so that a node must transmit data to its destination, generally the base station, with the help of other nodes in the network. The clustering protocols are mainly considering the cross layering techniques for designing energy efficient hierarchical wireless sensor networks (Sang, L., 2010). The sensor nodes that are belong to a cluster, send their sensed data to a node belong to their cluster called cluster head and then the cluster head eliminates the correlated data to reduce final data volume and send the aggregated data to the data sink. The clustering approaches can increase network longevity and improve energy efficiency by minimizing overall energy consumption and balancing energy consumption among the nodes during the network lifetime.

Existing routing protocols in the literature focus either on increasing the network lifetime or on addressing security issues, which consume a great deal of energy. None of them combines solutions to both challenges. Security is one of the major issues due to the unavailability of a physical line of defense. Therefore, WSN must work securely and be capable of detecting malicious activities. Attacks against WSN are grouped into two main categories: passive and active. Passive attackers are hidden and include such attacks as eavesdropping, node malfunctioning, node tampering / destruction, and traffic analysis. Active attacks directly affect network operations. They include denial of service (DoS), jamming, hole attacks (sink, warm, black), flooding, etc. Various security issues and their defenses are discussed in Related Work. There are two more kinds of attacks detected and eliminated from the network: sinkhole and Sybil. Sinkhole and Sybil attacks interrupt the

Corresponding Author: N.Vadivelan, Research Scholar, Department of Computer Science and Engineering, St.Peter's University, Avadi, Chennai, Tamilnadu, India.
E-mail: vadivelanresearch@gmail.com

normal nodes and intermediate nodes in the route as well as damage the data transmission process, thereby reducing the throughput of the network within the network area. Many existent studies provide detection and prevention mechanisms separately for individual attacks within the network.

In a perilous networking situation, adversaries may compromise link security or attempt to simulate legitimate nodes to perform malicious actions. Providing data encryption in communications is the first step in securing the network. Traditional public key schemes are not feasible because of processor and battery limitations in the sensor nodes, this makes symmetric cryptography very appealing. Recently, several key pre-distribution schemes (KPS) have been proposed to establish common keys, which are requisite for symmetric cryptosystems. In KPS, nodes are preloaded with key information prior to deployment. After deployment, nodes form secure communication links to neighbor nodes with a probability that is based on the key information shared between the two nodes.

Background Study:

There are various techniques are proposed and still in research for designing a routing protocol for wireless sensor networks. In the beginning of 21st century, wireless sensor networks are moving forward to finding practical abilities and finding new innovative applications (Habib, F., 2014). Wireless sensor networks are widely deployed, used and provide several wireless sensor networks such as WSS, WSAN, WISAN, WUSN, UWSN, WSIS, WDSS, WBSMN, SSN, UAVSN and IWSN (Labonte*et,* L., 2013). There are many issues and problems arise when the WSN application moves forward to large-scale common problems. Innovative solutions for recent and modern applications using WSN can be provided only in small-scale WSN applications. One-solution-many-problems for our most needed "SEE" applications are somewhat restricted (Wikipedia, 2014). Also, the traditional large homogeneous WSN answers cannot help to explore most scientific and industrial opportunities (Rashvand, H.F. and J.M. Alcaraz-Calero, 2012; Marques, L. and A. Casimiro, 2013).

In terms of energy, batteries are considered as the most important factor to be limited in WSS for SEE. The battery power can be saved using multi-state operations such as off, sleep, standby and use the power efficiency of the wireless spectrum (Rhee, S., 2004). Also, scaling-down the modulation (Schurgers, C., 2001), packet transmission by considering the properties of the sensor (Mukhopadhyay, S., 2004) are helping to save the energy. In WSN, most of the communication links are bidirectional. Due to the behavior and characteristics of the WSN applications, the communication may be in unidirectional (Sang, L., 2010, Ramasubramanian, V. and D. Mosse, 2008). In heterogeneous networks, communication in the opposite [reverse] direction is not possible (Wang, G., 2008). But, due to the ambient factors such as noise and interference it is necessary to lead the link as unidirectional. In recent applications, most of the MAC layer protocols are also using the bidirectional links. Routing protocols using MAC can utilize only bidirectional links for routing (Chen, B.B., 2009). Counting the hop length, a novel handshaking mechanism and ACK based unidirectional packet transmission are the key design parameters of a routing protocol to increase the lifetime of the network (Anil UfukBatmaz, 2014). Regular expressions (RegExes) are used to flexibly represent complex string patterns in many applications ranging from network intrusion detection and prevention systems(NIDPSs) ("Bro intrusion detection system," 2011) to compilers (Aho, A.V., 2007) and DNA multiple sequence alignment (Arslan, A.N., 2005; Chung, Y.S., 2007). In particular, NIDPSs Bro (2011) and Snort (2010) and Linux Application Level Packet Classifier (L7filter) (Levandoski, E., 2009) use RegExes to represent attack signatures or packet classifiers.

In this paper the contribution of our proposed approach is to discover an optimal route using Neighbor Selection in terms of Trust value, Energy and shortest distance. The route discovery concentrates mainly on optimizing the best trusted neighbor to obtain a hop-to-hop connection from source node to the destination node.

Proposed Approach:

In order to provide secured data transmission, the nodes information, behavior and data are analyzed, verified and validated as trusted. To do this, each parameter is represented as expressions and the combination of expression are assigned by a state and each state will change in a periodic interval. One of the main problems in wireless networking is to diagnose the nodes. Each node functioning in the network should maintain their correct information in terms of the status of the each component in the system like memory, battery level etc. In this paper, it is considered that the problem is to identify faulty nodes in the network using timed automata for representation. The fault analysis protocol is especially designed for WSN using FAT. Here, a new diagnosis algorithm is proposed to diagnose all the parameters related to the nodes in the network.

Assumptions and Notations:

In this paper, a new approach is created for the WSN for diagnosability using graph theory for denoting node connection topology in a timed variable linked to each node in the network. The networks is assumed by using the heartbeat messages where every node generates periodic messages and transmit to other nodes and it denotes that node is working. This scenario is called as time based graph,

Australian Journal of Basic and Applied Sciences, 9(20) June 2015, Pages: 90-95

and it is characterized by only one clock is reset into 0 for all edges. If the nodes starts and finishes within a finite time says that the node is correct and other nodes not finished within the finite is called as faulty nodes. A wireless sensor network described by the durational graph g and it is said to be $\delta-$diagnosable. If correct diagnosis is always provided then the number of faults does not exceed δ. The largest integer δ for which g is δ-diagnosable is called the diagnosability of g.

A direct graph can be identified as a tuple $g = \{Q, Q_0, E, \Psi, \eta, Inv, G, t\}$, where t denotes the time, and Q denotes the discrete space, Q_0 denote the initial condition of the discrete space $\eta: E \to \Psi$ represents the resultant function.

In this paper, it is assumed that:

1. Entire node in the network has a unique ID and encoded.
2. Every node knows its ID and its nearest neighbors also.
3. Fault can be detected during the node diagnosis process.
4. Communication among the nodes are symmetric and as a connected graph.

Given an automation A, let $P_c \subseteq P$ is a discrete state, representing a fault model in A: P_c is known as faulty set. If the nodes are diagnosable for some finite δ, then the set of all properties shows that it is very interesting to calculate the least value δ_m for which A is δ_m-diagnosable: given A, the following statements holds:

1. If P_c is a δ_m-diagnosable, then it is $\delta^* -$diagnosable for all $\delta^* \leq \delta$.
2. If P_c is not δ_m-diagnosable, then it is not $\delta^* -$diagnosable for all $\delta^* \leq \delta$.

Network Model:

In this paper, the network is comprised of N sensor nodes deployed in random locations, communicating through radio transceivers. Entire nodes in the network are homogeneous and functioning with a limited energy supply. All the nodes are behaving as source node as well as sink node. When the node is used as a potential sink it can be used by an external operator to fetch the information gathered by the network. All the sensor nodes are activated in two states such as faulty and faultless. Faulty nodes are incapable to converse with the rest of the system, either due to a crash or to battery depletion. This means that faults are permanent, i.e. nodes remain faulty until they are repaired and/or replaced. This paper, the communication among the nodes may be unicast or multicast. Various fault-diagnoses are applied on the nodes, such as node-crash-fault, $\delta - fault$ and node-misbehavior-fault. Node-crash-fault is diagnosed using heartbeat-based diagnosis and other faults are diagnosed by $\delta - faulty$ execution and verifying expressions of node's. It is assumed that data transmission from node u is Omni directional. Due to this, the message sent by node u can be received by of its neighbor nodes within the transmission range. The neighbor of u is represented by N(u). The entire network topology is described as a graph G =(V, E) named as communication graph. V is the set of nodes and E is the set of edges connecting nodes in G. For any $u, v \in V$, directed edge $(u, v) \in E$ if and only if $v \in N(u)$. The set of all faulty sensor nodes are denoted as F, with $|F| = f \geq 0$.

MATERIAL & METHODS

In this paper, our proposed algorithm is initiated by a unique faulty-less node, henceforth it is called as gateway nodes and they will become cluster heads used for external operations. Two types of messages are exchanged during execution: 'Hand-shaking' messages and 'diagnostic' messages. The 'Hand-shaking' messages sent by cluster-head u is a pair of identifiers *(u,v),* where *v* is the identifier of the node from which u received the first 'hand-shaking' message (u itself, if u is the cluster-head). 'Hand-shaking' messages are used to detect faulty nodes. The diagnostic message sent by node u is a pair (u, Qc), where Qc is the set of the identifiers of the nodes currently diagnosed as faulty by node u. Diagnostic messages are used to propagate the identities of faulty sensors throughout the network. The overall functionality of our proposed approach is written in the form Pseudo code for implementing and verifying the performance.

Pseudo Code:

Diagnose ( )

{

1. Construct the network G with N number of nodes, connected with E number of edges.
2. Initialize all the relevant parameters such as Initial Energy, energy for Tx and energy for Rx, R [radio transmission range] etc.
3. Let BS be the base station
4. Terminate= false;
5. The entire nodes handshake with BS and among themselves.
6. Arrange all the nodes according to their energy value in descending order. This helps to find out the gateway nodes.
7. A request sent to gateway nodes and gateway node ping to node back.
8. Node-i chosen as initial node and it broadcast a request to all its neighbors. A time interval is maintained for the request messages and time-out is set for diagnosing the nodes. Whichever the node gets time-out message is treated as faulty node.
9. If a node N_i within the range of P_c, then N_i sends a "Hi" message to node N_j, used to compare with N_i for further verification. During the verification it receives (N_i, P_c).

$\forall\rho \in \cup F\delta^*, \forall\rho^1 \in \mathcal{L}\backslash\cup F\delta^*, P(\rho) \neq P(\rho^1))$ derived from the following condition.
$\forall\rho \in \cup F\delta^*, \forall\rho^1 \in \mathcal{L}\backslash\cup F\delta^*, P(\rho) \neq P(\rho^1))$. where P_c is defined as observable for a system H if it is possible to immediately detect using the observable output whether the current discrete state is visiting P_c.
10. If match occurs N_i diagnosis P_i to be faulty-less. If all the fault neighbors sent their diagnosis, the CH's diagnosis is complete and it sent to all the faulty-less nodes in the network by broadcasting a message. Else, N_i diagnosis P_i to be faulty, if the node is out of range or did not reply within timeout condition.
11. Otherwise, node N_i diagnosis node P_i as faulty If (P_c=Number(CH)) then diagnosis complete.
12. Terminate=true;
13. End if;
14. }

Analysis For Proposed Algorithm:

In this section, we verify the proposed algorithm. A WSN can be modeled by a group of N nodes a set of edges connecting the nodes within a radio link. The number of nodes deployed in the network is 10, 20 30, 40 and 50. $\sqrt{N}$ number of nodes are elected as CH. A connection is created among the nodes and the CH according to the radio communication range. A timer is set for communication like time between sending and receiving a message among the nodes as well as between the CH and time for CH to CH message passing. If the node is out of range or if the node is not reply within a time interval it is considered as crash faulty node in the network.

Simulation and Result:

These above procedures are simulated in Network Simulator2 software and verify the performance of the proposed approach. Also it is evaluated the latency of the message transmission complexity. It is considered an abstract procedure for verifying the diagnosability algorithm in a network. The minimum and maximum message delays Δ_{send_min} and Δ_{send_max} are the minimum and maximum times, respectively among the last bits of message are being sent into the network and message being completely delivered at a working neighbor node. In this paper, each node in the network is a entity. In the entire network, the total number of messages passed within a time interval and number of faulty nodes, faulty-less nodes with delay are diagnosed by our diagnosis algorithm.

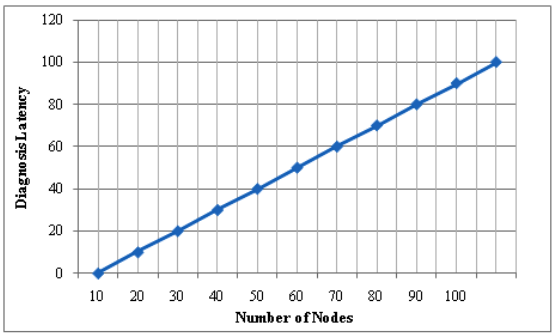

Fig. 1: Diagnosis Latency of Our Proposed Algorithm.

By assigning a set of determined delay the algorithm is executed by node simulation. From the above figure-1, it is clear that the message complexity in time is presented. The total number of messages are passed to diagnosis the faulty nodes and faulty-less nodes in the network. In case, if the network increases then the number of messages is increased. Here we use the time event for computing the time delay of messages. Within the time interval only the nodes can send and receive the messages, if not those nodes are found as faulty nodes. Once the node is found as faulty node then that information passed to CH and the CH passed to other CH and broadcast globally. Figure-1 shows that the diagnosis latency is depend on the number of nodes in the network. It shows that the number of node increases the diagnosis time gets increases. In our proposed algorithm the initial time set as 0.02.

In our simulation, each node assigned with certain number of messages to be passed within a time interval. There are 100 number nodes are deployed in 10 rounds such as 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 nodes in the network for simulation based performance analysis. Figure-2 shows that the number of messages passed within the time interval in terms of number of nodes. For 10 numbers of nodes 86 messages are passed within time interval, 20 number of nodes 167 messages are passed within time interval, 30 number of nodes 167 messages are passed within time interval, 40 number of nodes 167 messages are passed within time interval, 50 number of nodes 167 messages are passed within time interval, 60 number of nodes 167 messages are passed within time interval, 70 number of nodes 167 messages are passed within time interval, 80 number of nodes 167 messages are

passed within time interval, 90 number of nodes 167 messages are passed within time interval and 100 number of nodes 167 messages are passed within time interval. From figure-2, it is computed that 1.05% of messages only not passed within the time interval assigned.

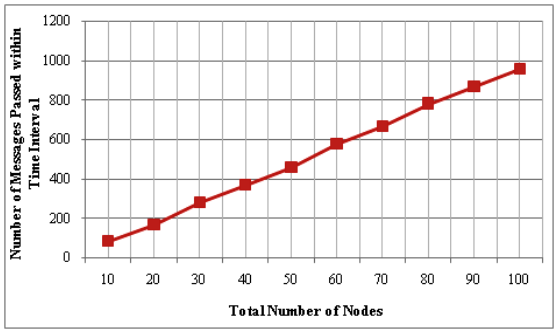

Fig. 2: Number of Node vs. Messages Passed within Time Interval.

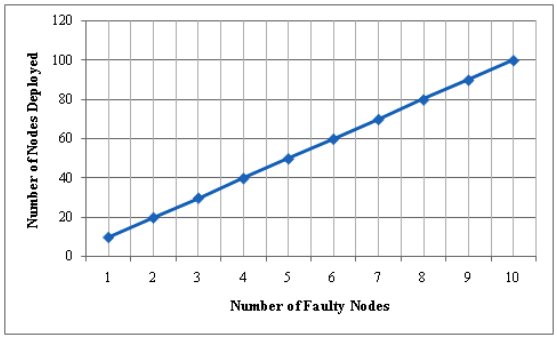

Fig. 3: Number of Nodes vs. Number Faulty Nodes .

From Figure-3, it shows that only less number of nodes is found as faulty nodes in the network. For 10 numbers of nodes there is no node found as faulty node. For 20, 30 nodes only one node is found as faulty node, for 40, 50, 60 and 70 nodes there are 2 nodes are found as faulty nodes and for 80, 90 and 100 number of nodes the number of faulty nodes found are 3.

Conclusion:

In this paper, it is assumed that the gateway nodes are treated as intermediate or interface nodes in the network which can accept the ping control inputs and replies within a time delay. It is stated that the important condition for being able to detect the faulty nodes using our proposed algorithm using the diagnosability model. From the simulation results figure-1 to figure-3, it is clear that our proposed approach using automata based diagnosability verification is out performs in terms of latency and detecting faulty nodes and computing delay. We simulated our proposed approach algorithm using NS2. The diagnosis parameters are diagnosis latency, message complexity and number of faulty nodes in the network. From this is concluded that if the network size increases the diagnosis latency and message complexity of the network is also gets increased.

REFERENCE

Inhye Park, Hyungkeun Lee and Seokjoong Kang, 2014 "RIX-MAC: An Energy-Efficient Receiver-Initiated Wakeup MAC Protocol for WSNs", KSII Transactions On Internet And Information Systems, 8-5.

NityanandaSarma and Sukumar Nandi, 2010. " A Multipath QoS Routing with Route Stability for Mobile Ad-Hoc Networks", IEEE –Technical Review.

Habib, F., Rashvand, Ali Abedi, Jose M. Alcaraz-Calero, Paul D. Mitchell and Subhas Chandra Mukhopadhyay, 2014. "Wireless Sensor Systems for Space and Extreme Environments: A Review", IEEE Sensors Journal, 14-11.

Labonte*et*, L., 2013. "Wireless sensor and actuator networks with delayed noisy feedback (WiSAN)," in Proceeding of IEEE International Conference on Wireless Space Extreme Environment(WiSEE*)*, Baltimore, MD, USA, 1-5.

Wikipedia, 2014. Wireless Sensor Network,[Online]Available:http://en.wikipedia.org/wiki/Wireless_Sensor_Network.

Rashvand, H.F. and J.M. Alcaraz-Calero, 2012. Distributed Sensor Systems: Practice and Applications. London, U.K.: Wiley.

Marques, L. and A. Casimiro, 2013. "Fighting uncertainty in highly dynamic wireless sensor networks with probabilistic models," Proceeding of 32nd International Symposium Reliable Distributed System, 31–40.

Rhee, S., D. Seetharam and S. Liu, 2004. "Techniques for minimizing power consumption in low data-rate wireless sensor networks," Proceeding of IEEE Wireless Communications and Networking Conference (WCNC), Atlanta, GA, USA, 1727–1731.

Schurgers, C., O. Aberthorne and M.B. Srivastava, 2001. "Modulation scaling for energy aware communication systems," Proceeding of ACM International Symposium Low Power Electronics and Design, Huntington Beach, CA, USA, 96–99.

Mukhopadhyay, S., D. Panigrahi and S. Dey, 2004. "Data aware, low cost error correction for wireless sensor networks," Proceeding of IEEE Wireless Communications and Networking Conference (WCNC), Atlanta, GA, USA, 2492–2497.

Ramasubramanian, V. and D. Mosse, 2008. "BRA: A bidirectional routing abstraction for asymmetric mobile ad hoc networks," IEEE/ACM Transaction on Networking., 16(1): 116–129.

Sang, L., A. Arora and H. Zhang, 2010. "On link asymmetry and one-way estimation in wireless sensor networks," ACM Transaction on Sensor Networks, 6(2): 12:1–12:25, Art. ID 12.

Wang, G., D. Turgut, L. Bölöni, Y. Ji and D.C. Marinescu, 2008. "A MAC layer protocol for wireless networks with asymmetric links," Ad Hoc Networks Journal, 6(3): 424–440.

Chen, B.B., S. Hao, M. Zhang, M.C. Chan and A.L. Ananda, 2009. "DEAL: Discover and exploit asymmetric links in dense wireless sensor networks," Proceeding of 6th Annual IEEE Communication Society Conference on Sensor, Mesh, Ad Hoc Communication and Networking (SECON), 1–9.

Anil UfukBatmaz, HuseyinUgurYildiz and BulentTavli, 2014. "Role of Unidirectionality and Reverse Path Length on Wireless Sensor Network Lifetime", IEEE Sensors Journal, 14-11.

"Bro intrusion detection system," 2011 [Online]. Available: http://www.bro-ids.org

"Snort network intrusion detection system," Source fire, Columbia, M.D., 2010 [Online]. Available: http://www.snort.org.

Aho, A.V., M.S. Lam, R. Sethi and J.D. Ullman, *2007*. Compilers: Principles, Techniques, and Tools, 2nd ed. Reading, MA: Addison-Wesley.

Arslan, A.N., 2005. "Multiple sequence alignment containing a sequence of regular expressions," Proceeding of IEEE CIBCB, 1–7.

Chung, Y.S., W.H. Lee, C.Y. Tang and C.L. Lu, 2007. "RE-MuSiC:Atool for multiple sequence alignment with regular expression constraints," Nucleic Acids Res., 35: W639–W644.

Levandoski, E., Sommer and M. Strait, 2009. "Application layer packet classifer for Linux," [Online]. Available: l7-filter.sourceforge.net.