Anomaly Detection in Wireless Sensor Networks: A Proposed Framework

https://doi.org/10.3991/ijim.v14i10.14261

Dina M. Ibrahim (✉), Nada M. Alruhaily
Qassim University, Buraydah, Saudi Arabia
d.hussein@qu.edu.sa

Abstract—With the rise of IOT devices and the systems connected to the internet, there is, accordingly, an ever-increasing number of network attacks (e.g. in DOS, DDOS attacks). A very significant research problem related to identifying Wireless Sensor Networks (WSN) attacks and the analysis of the sensor data is the detection of the relevant anomalies. In this paper, we propose a framework for intrusion detection system in WSN. The first two levels are located inside the WSN, one of them is between sensor nodes and the second is between the cluster heads. While the third level located on the cloud, and represented by the base stations. In the first level, which we called light mode, we simulated an intrusion traffic by generating data packets based on TCPDUMP data, which contain intrusion packets, our work, is done by using WSN technology. We used OPNET simulation for generating the traffic because it allows us to collect intrusion detection data in order to measure the network performance and efficiency of the simulated network scenarios. Finally, we report the experimental results by mimicking a Denial-of-Service (DOS) attack.

Keywords—Anomaly detection, wireless sensor network, DoS attack, OPNET simulator

1 Introduction

A wireless sensor network (WSN) is a common network architecture made up of a set of autonomous devices and sensor nodes for gathering data from the adjacent environment. Examples of collected data sources are humidity sensor nodes, temperature sensor devices, power, light, etc. The need of wireless sensor networks is growing continuously, because of the enormous improvement of technology [1]. Simultaneously, effective administration techniques are needed for dealing with complex networks and with the disparities of sensor data [2,3]. Wireless sensor networks are naturally associated with cloud services over the Internet. The storage and computing infrastructures are provided by the Cloud platforms that are necessary for archiving, analyzing, and processing the huge amount of data produced by sensors [4,5].

Anomaly detection in homogeneous WSNs received much attention in the literature. Most of the methods concerning with anomaly detection are devoted to the analysis of data streams produced by any single device [6,7]. In this issue, any node device is analyzed, by using many different methods, to realize whether or not an anomaly has been detected. These methods and techniques are generally based on composite mathematical analysis or statistical or numerical methods applied on data streams [8,9], which is appropriate for the specific arithmetic characteristics of the sensor data type. Consequently, WSNs applied such techniques in order to sense diverse type of features/parameters, and including many sensors is not simple. Authors in [10], proposed Short-long term anomaly detection in wireless sensor networks based on machine learning and multi-parameterized edit distance, their method is performed by applying the analysis of edge and cloud on real data, which has been developed inside residential building and then deformed with a set of fake impairments.

2 The Proposed IDS Framework

In this paper, we propose a framework for intrusion detection in WSN; the WSN designed based upon three elements: the sensor nodes, the cluster heads, and the base stations, as illustrated in Fig. 1. The framework consists of three levels: the first level is the sensor node level, that we can call it the short-term level where the nodes monitored by the cluster head. If any of the sensors detect any type of anomaly, it can forward this thread to the correspondence Cluster head. The second level is the cluster head level, that we call it the medium-term level, it deals only with the cluster heads inside the network. In this level, the base stations monitor the cluster heads where the last can send the detected anomaly to the base station. We can combine the monitoring processes in the sensor nodes and the cluster head to constitute an edge-based method, as represented in Fig. 1.

The third and last level is the base station level, which we call it the long-term level and the monitoring process here is done using a cloud-based method. In this level the monitoring process between the base stations is done outside the WSN (on the cloud). The cloud-based method uses the historical data in addition to the information sent by the base stations to take a suitable decision. Inside the WSN; Edge-based Method is effective on identifying short term anomalies while In the Cloud: Cloud-based method: is more accurate on identifying long term anomalies that uses the cloud storage sensed data, the historical analysis, and the new sensed data from the base stations to detect the long-term anomalies.

Simulation and modelling are important approaches in the development and evaluation of the systems in terms of time and costs. Simulations are used to show the expected behavior of the system under different conditions [11]. There are many different possible platforms for simulation and testing anomaly detection framework on WSNs. The current WSNs simulators allow users to measure different features by varying different parameters. Example of features that could be measured are: simulation scenarios, global behavior, energy effectiveness, and fault tolerance. Each of them can be measured with different simulation programs, however, the most popularly used simulators for WSNs are: OPNET, NS-2, SensorSim, GloMoSim, OMNET++, and JavaSim [12]. In addition to the possibility of using hybrid simulators that can use both real devices to sense the data integrated with suitable simulator of produce cheaper software development system that can increase the reliability on the results [13].

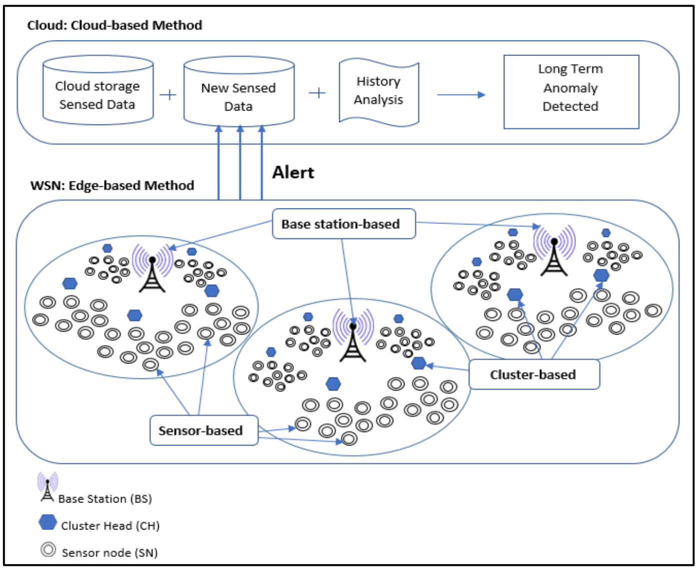


Fig. 1. The proposed IDS framework in Wireless Sensor Network

3 Simulation Model Using OPNET

We are using OPNET Modeler for this paper because of the numerous benefits it offers. OPNET provides a Graphical User Interface (GUI) for the designed model or topology, which allows a realistic simulation of networks; it has also a performance display module and data collection [14]. Another benefit of using OPNET Modeler is that it has been used widely by network researchers, and there is a wide confidence in the validity of the produced results. OPNET permits an accurate analysis of performance measures and an effective detection of network intrusions.

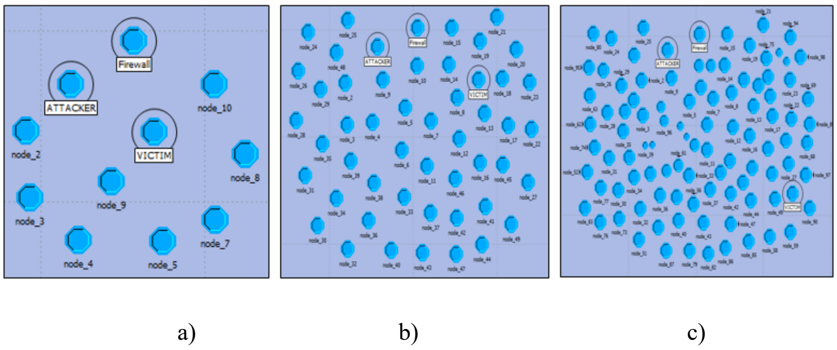
The network traffic source comes from the MIT/Lincoln Lab TCPDUMP files. It contains various simulated network attacks, including DoSNuke attack packets, which has been used in this paper for testing the proposed model [15]; DoSNuke attack is a Denial of Service (DoS) attack, which exploits a known vulnerability in Windows NT operating system. We could use, alternatively, software tools such as Nmap [16] to generate attacks while running a network sniffer such as Ethereal [17] to capture the network traffic, with the attack and network sniffing activities all occurring in a controlled lab environment. The captured Ethereal file, which includes the attack data and normal user data if desired, can be used in our intrusion simulation experiment. For both the TCPDUMP and Ethereal files, we use pre-processing tools to extract information of the traffic, which is very important for OPNET simulation. Our tools can analyze the TCPDUMP and Ethereal files, extracting the flag information, the data packet headers, the time distribution, and the packet payload. OPNET software can also provide an Application Characterization Environment module (ACE) that can be used to include packet traces into simulation, supporting packet formats of various sources as well as TCPDUMP files [18].

Before building the simulated network, we need first to pre-process the TCPDUMP file (or Ethereal file) to extract the relevant information, including the packet inter-arrival times, which are saved as a list of double-type values. In addition to the time duration, which is the time difference between the first packet and the last packet of the traffic source, and a list of the distinct IP addresses in the traffic source.

3.1 Building the network model

To get better results from our model; we build three scenarios for our WSN. The first scenario consists of a coordinator that represent the cluster head in our framework, a router that acts as a firewall, and 10 sensor nodes (or end devices), as shown in Fig. 2 (a), it represents the OPNET model for simulating the DoSNuke attack, sensor Node 1 is the ATTACKER while sensor Node 6 is the VICTIM. The packets are extracted from the traffic source file. Once a packet is ready to be sent from the source to the destination (according to the predefined traffic scenario) it can be transmitted without delay and also the traffic flow is consistent.

The Second scenario for simulating the DoSNuke attack is shown in Fig. 2 (b). It consists of 50 sensor nodes; sensor Node 1 is the ATTACKER while sensor Node 20 is the VICTIM. Figure 2 (c) shows the third scenario which have 100 sensor nodes with Node 1 is the ATTACKER and Node 6 is the VICTIM of the DoSNuke attack. There is a firewall node between the victim and the coordinator, which we use to capture suspicious data packets to, or from the victim using the DoSNuke attack's signature.



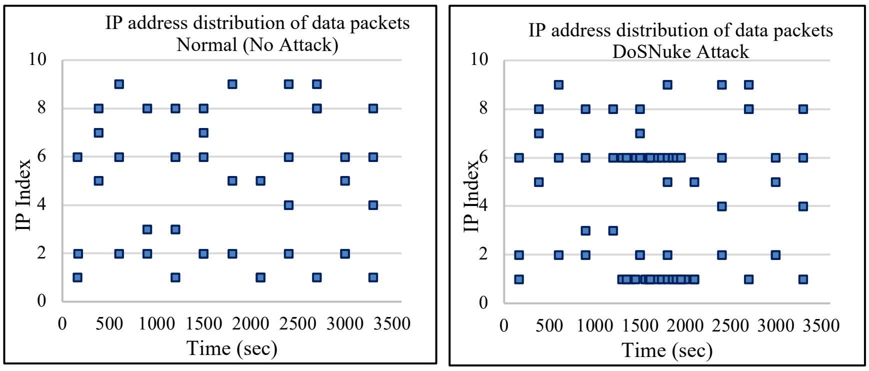
a) b) c)

Fig. 2. The network model in three scenarios. 10 nodes, 50 nodes, and 100 nodes

During the configuration process for the OPNET model, we build a generator module to use a predefined interval times for the script file that is responsible for generating the packets. This file is the output of a pre-processing step of the source traffic.

3.2 Analysis of simulation results

The source traffic data file for the DoSNuke attack comes from the MIT/Lincoln Lab, DARPA intrusion detection evaluation data set, outside TCPDUMP dataset, 1999/Week 4/Wednesday. The simulated network normally collected data twenty-two hours a day. The tcpslice program was used to examine the outside TCPDUMP data files, and the actual times of the first and last packet were extracted. [19]. We set up several statistical measures in OPNET to study the performance of the intrusion simulation. One of these measures is the IP address distributions of the data packets during the entire simulation, as explained in Figures 3, 4, and 5 that represent the three scenarios; 10 Nodes, 50 Nodes, and 100 Nodes, respectively.

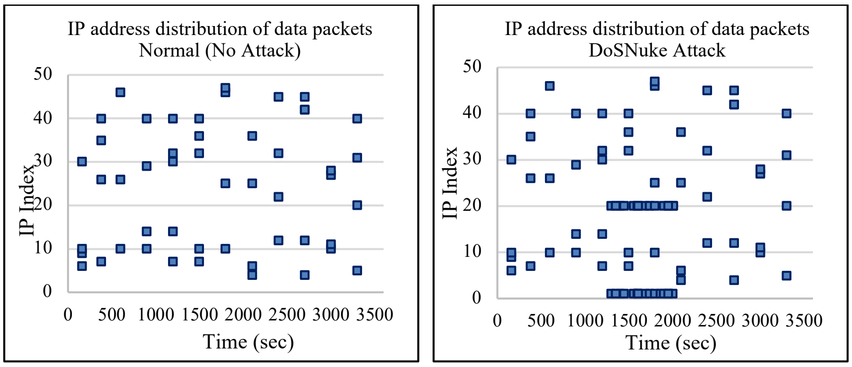


a)

b)

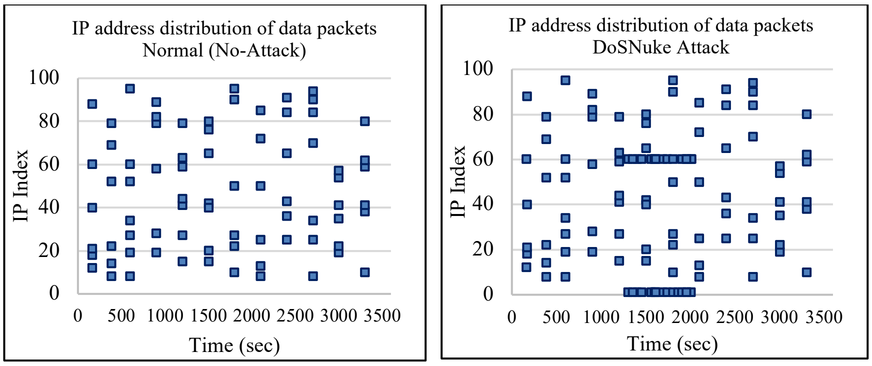
Fig. 3. IP address distribution of data packets (10 Nodes scenario) (a) Normal mode and (b) DoSNuke Attack

The figures clearly demonstrate that in the normal mode without attack the accesses to the IP addresses is consecutive while on the DoSNuke attack mode there is a clear crowded access on the IP for nodes 6, 20, 60 as in figures 3(b), 4(b), and 5(b), respectively, in range approximately from 1200 sec to 2200 secs.



a) b)

Fig. 4. IP address distribution of data packets (50 Nodes scenario) (a) Normal mode and (b) DoSNuke Attack



a) b)

Fig. 5. IP address distribution of data packets (100 Nodes scenario) (a) Normal mode and (b) DoSNuke Attack

The average rates of data packets captured by the firewall router, which means the occurrences of the packets and the times of their arrivals is illustrated in Fig. 6. The figure shows the firewall traffic for the sent and received data with normal mode (No-attack) and with DoSNuke attack. It is clearly noticed in the figure that the traffic has high rates during the period 1200 sec and 2200 sec, which means at the same periods appeared in figures 3, 4, and 5. These results demonstrates the occurrences of the DoSNuke attack captured synchronously with the firewall.

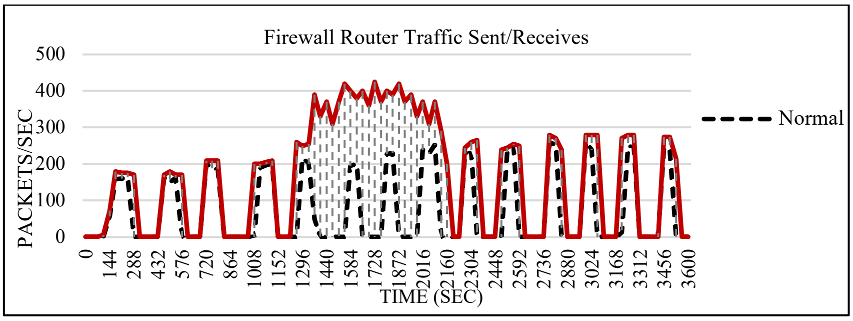


Fig. 6. The firewall router traffic sent/received with Normal (No-attack) and DoSNuke Attack

In Fig. 7, we collected the overall network traffics statistics in the normal mode and with DoSNuke attack mode. The figure shows that the traffics in normal mode is stable while in the DoSNuke attack the traffic increases in synchronous with the same period of time appeared in the three scenarios represented in figures 3, 4, and 5.

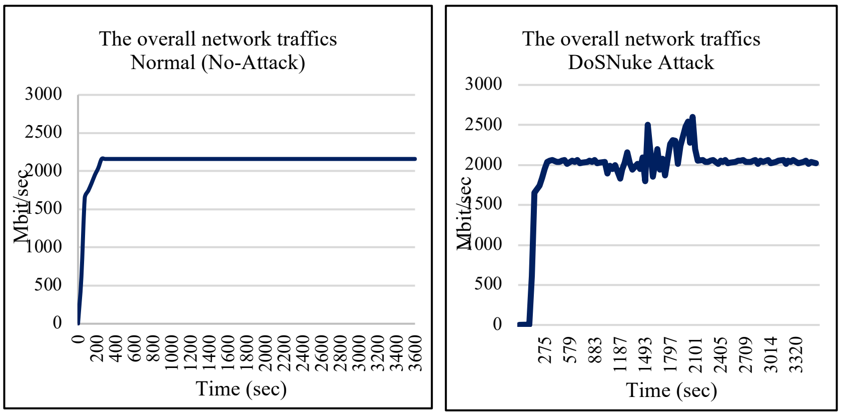


Fig. 7. The overall network traffic with Normal (No-attack) and DoSNuke Attack

4 Conclusion

In this paper, we propose a framework for intrusion detection system in WSN. The first two levels are located inside the Wireless Sensor Networks (WSN), one of them is between sensor nodes and the second between the cluster heads. While the third level located on the cloud and represented by the base stations. In the first level, which we called light mode, we simulated an intrusion traffic by generating data packets based on TCPDUMP data, which contain intrusion packets. Then, we reported experimental results of network intrusion simulation using previously captured TCPDUMP data as the traffic sources. Our work demonstrated several aspects using OPNET Modeler simulator for detecting intrusions by displaying and identifying patterns of the IP address distributions of the data packets during the entire simulation using three different scenarios and the average rates of data packets captured by the firewall router.

For future work, we plan to complete the evaluation of the second and third levels of our proposed framework in WSN with different intrusion detection aspects.

5 References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, "Wireless sensor networks: a survey," International Journal of Computer Network, vol. 38, no. 4, pp. 393-422, 2002. https://doi.org/10.1016/s1389-1286(01)00302-4

[2] G. Fortino, R. Giannantonio, R. Gravina, P. Kuryloski, R. Jafari, Enabling effective programming and flexible management of efficient body sensor network applications, IEEE Trans. Hum. Mach. Syst., vol. 3, no. 1, pp. 115–133, 2013. https://doi.org/10.1109/tsmcc.2012.2215852

[3] G. Fortino, A. Guerrieri, G.M. O'Hare, A. Ruzzelli, "A flexible building management framework based on wireless sensor and actuator networks," International Journal of Network Computer Application, vol. 35, no. 6, pp. 1934–1952, 2012. https://doi.org/10.1016/j.jnca.2012.07.016

[4] J.A. Stankovic, "When sensor and actuator networks cover the world," ETRI Journal, vol. 30, no. 5, pp. 627–633, 2008. https://doi.org/10.4218/etrij.08.1308.0099

[5] H.H. Bosman, A. Liotta, G. Iacca, H. Wortche, "Anomaly detection in sensor systems using lightweight machine learning," in: Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on, IEEE, pp. 7–13, 2013. https://doi.org/10.1109/smc.2013.9

[6] Y. Zhang, J. Jiang, "Bibliographical review on reconfigurable fault-tolerant control systems," Annual Revision Control, vol. 32, no. 2, pp. 229–252, 2008. https://doi.org/10.1016/j.arcontrol.2008.03.008

[7] S. Ahmad, A. Lavin, S. Purdy, Z. Agha, "Unsupervised real-time anomaly detection for streaming data," International Journal of Neurocomputing, vol. 26, no. 2, pp. 134–147, 2017. https://doi.org/10.1016/j.neucom.2017.04.070

[8] Y. Yao, A. Sharma, L. Golubchik, R. Govindan, "Online anomaly detection for sensor systems: a simple and efficient approach," International Journal of Performance Evaluation, vol. 67, no. 11, pp. 1059–1075, 2010. https://doi.org/10.1016/j.peva.2010.08.018

[9] M. Xie, S. Han, B. Tian, S. Parvin, "Anomaly detection in wireless sensor networks: a survey," International Journal of Network Computation Applications, vol. 34, no. 4, pp. 1302–1325, 2011. https://doi.org/10.1016/j.jnca.2011.03.004

[10] F. Cauteruccio, G. Fortino, A. Guerrieri, A. Liotta D. Mocanu, C. Perra, G. Terracina, and M. Vega, "Short-long term anomaly detection in wireless sensor networks based on machine learning and multi-parameterized edit distance," International Journal of Information Fusion, vol 52, no. 2019, pp. 13-30, 2019. https://doi.org/10.1016/j.inffus.2018.11.010

[11] Mosad Alkhathami, Lubna Alazzawi, and Ali Elkateeb. "Border Surveillance and Intrusion Detection using Wireless Sensor Networks," International Journal of Advances in Engineering & Technology (IJEAT), vol. 8, no. 2, pp. 17-29, 2015. https://doi.org/10.17577/ijertv4is040150

[12] J. Rout, "A Novel Clustering Protocol in Wireless Sensor Network," 2019 International Conference on Applied Machine Learning (ICAML), Bhubaneswar, India, 2019, pp. 258-261. https://doi.org/10.1109/icaml48257.2019.00054

[13] OPNET Technologies, OPNET Modeler Product Documentation Release 15.0, Inc., Washington DC, 2009.

[14] S. Saginbekov and C. Shakenov, "Hybrid simulators for wireless sensor networks," 2016 IEEE Conference on Wireless Sensors (ICWiSE), Langkawi, 2016, pp. 59-65 https://doi.org/10.1109/icwise.2016.8188543

[15] DARPA Intrusion Detection Evaluation dataset, at https://www.ll.mit.edu/r-d/datasets/1999-darpa-intrusion-detection-evaluation-dataset, [Last Accessed 20/02/2020].

[16] NMAP, at http://www.insecure.org., [Last Accessed 20/02/2020].

[17] Gerald Combs, Ethereal – Network Protocol Analyzer, http://www.ethereal.com/, [Last Accessed 20/02/2020].

[18] S. Razak, M. Zhou, and S. Lang, "Network intrusion simulation using OPNET." International Journal of Computer Science, pp. 1-5, 2002.

[19] 1999 DARPA Intrusion Detection Evaluation Data Set, Training Week 4, outside TCPDUMP data at Wednesday https://archive.ll.mit.edu/ideval/data/1999/testing/week4/index.html, [Last Accessed 20/02/2020].

6 Authors

Dina M. Ibrahim Assistant Professor at Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia from September 2015 till now. In addition to, Dina works as Lecturer at Computers and Control Engineering Department-Faculty of Engineering-Tanta University-Egypt. She was born in United Arab of Emarat, her B.Sc., M.Sc. and Ph.D. degrees taken from Computers and Control Engineering Department-Faculty of Engineering, Tanta University at 2002, 2008, and 2014, respectively. Dina Works as consultant Engineer, then a Database administrator, and finally acts as a Vice Manager on Management Information Systems (MIS) Project, Tanta University, Egypt, from 2008 until 2014. (email: d.hussein@qu.edu.sa).

Nada M. Alruhaily works as an Assistant Professor at Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia She is a Cybersecurity consultant and researcher with a special interest in Malware analysis & detection and related Malware attacks. She got her PhD on 2018 in Malware Detection Mechanisms from the University of Birmingham - UK (email: nrhiely@qu.edu.sa).

Article submitted 2020-03-24. Resubmitted 2020-04-27. Final acceptance 2020-04-27. Final version published as submitted by the authors.