SECURITY AND COMMUNICATION NETWORKS
Security Comm. Networks 2016; **9:**2744–2751
Published online 11 February 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/sec.1181

SPECIAL ISSUE PAPER

An intrusion detection method for wireless sensor network based on mathematical morphology

Yanwen Wang[1], Xiaoling Wu[1,2,3,4]* and Hainan Chen[1,5]

[1] Guangzhou Institute of Advanced Technology, CAS, No. 1121, Haibin Rd, Nansha District, Guangzhou, 511458, China
[2] Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis, Guangdong University of Petrochemical Technology, Maoming, 525000, China
[3] Information Technology Research Base of Civil Aviation Administration of China, Civil Aviation University of China, Tianjin, 300300, China
[4] Shenzhen Institutes of Advanced Technology, CAS, No. 1068, Xueyuan Rd, Shenzhen University Town, Nanshan District, Shenzhen, 518055, China
[5] Guangdong University of Technology, No.100, Huanchengxi Road, University Town, Guangzhou, 510006, China

ABSTRACT

Security issue in Internet of Things (IoTs) has long been the topic of extensive research in the last decade. Data encryption and authentication are the most common two methods to address the security issues in IoTs. However, these efforts are ineffective in detecting the diverse malicious attacks, especially in intrusion detection. Comparatively, very few attentions have been paid for detecting intrusive nodes in IoTs research. Therefore, in this paper, we derive an innovative method called granulometric size distribution (GSD) method based on mathematical morphology for detecting malicious attack in IoTs, such as intrusion detection. We successfully generate GSD clusters to directly monitor the number of active nodes in a wireless sensor network because the GSD curves are similar when the number of active nodes in a wireless sensor network is fixed. Link Quality Indicator data of each node are utilized as the network parameters in this method. The results show the effectiveness in intrusion detection. Copyright © 2015 John Wiley & Sons, Ltd.

KEYWORDS
internet of things (IoTs); intrusion detection; mathematical morphology (MM); GSD; LQI

***Correspondence**
Xiaoling Wu, Guangzhou Institute of Advanced Technology, CAS, No. 1121, Haibin Rd, Nansha District, Guangzhou, 511458, China.
E-mail: xl.wu@giat.ac.cn

1. INTRODUCTION

Miniature intelligent sensors can be deployed in harsh environment instead of manual manipulation. These sensors are capable of automatically monitoring diverse environment metrics by collecting and processing the data in real time and transmitting the corresponding data to other sensor nodes or a base station. The rise of wireless sensor networks (WSNs) has paved the way for the construction of Internet of Things (IoTs) systems in various social core departments [1]. Sensors embedded in IoTs devices are distinguished by being battery embedded, self-organized and highly reliable and available, which attract increasing attention in recent research.

However, in IoTs, data privacy and security are one of the indispensable issues that should be carefully considered when operating the entire networks. Because of the reason that IoTs are invented for collecting and transmitting information instead of manual labors, potential security issues might exist during the data transmission. For example, in a predeployed IoTs system, one or even more than one “spy nodes" can be abruptly placed into it. Data that are transmitted among devices might be duplicated and eavesdropped through these “spy nodes" and then transmitted to another IoTs system, which results in great property loss. Because this information interception behavior is automatically and individually accomplished by the wireless sensor devices in IoTs without any personnel, it is hard to be detected, especially in a hostile environment.

The rest of this paper is organized as follows. In Section 2, Link quality indicator (LQI) properties and mathematical morphology (MM) technique will be briefly introduced. In Section 3, some related work will be

Copyright © 2015 John Wiley & Sons, Ltd.

presented. In Section 4, the proposed Granulometric Size Distribution (GSD) method will be described in detail with mathematical expression. The results will be shown in Section 5, and finally, we conclude our method in Section 6.

2. RESEARCH BACKGROUND

2.1. The utilization of link quality indicator

Link quality indicator data are used as a major metric in this paper. LQI is frequently used in wireless networking [2,3] to indicate how strong the communications link is. In other words, LQI represents the quality reception of a packet in a communication link [4]. The LQI is a specific parameter to estimate how easily a received signal can be demodulated based on the analysis of error rate, and it can be directly read from the received packets structure. For the most widespread radio (CC2420), the value of LQI is ranging from 50 to 110 [4], which can be extracted from first eight symbols of the received packet. One of the most fulfilling reasons that we use LQI data in this paper is that LQI is able to assess the channel with a high degree of reliability without acquiring too much resources. In this paper, we distinguish different patterns when different numbers of nodes are communicating in the WSN through the application of GSD method to deal with the LQI data. Figure 1 shows a simple example of LQI collected from the same node when different numbers of devices are communicating in an IoT system. As we can see in Figure 1, the original LQI data for the same device when two, three and four devices are communicating are shown in Figure 1(a)–(c), respectively. Figure 1(d) combines all LQI of these three scenarios. Obviously, it is very hard to detect a “spy node" only by distinguishing the original LQI data because they might overlap each other.

2.2. Mathematical morphology

Mathematical morphology was first introduced in the 1960s by Matheron and Serra [5]. MM consists of theory and techniques for the analysis of spatial structures such as shape and size of objects. It is based on Set Theory, Integral Geometry and Lattice Algebra and has been widely used in Image Processing. The four basic important operators in MM are dilation, erosion, opening and closing. To accomplish these four operators, Structuring Element (SE) is used. SEs are kind of basic shapes, which are simple and predefined, such as squares, rectangles, triangles and circles. The key point of processing a binary image is to use SE to probe (dilate or erode) a target object.

It is understandable that erosion operator erodes (shrinks) the target objects. Mathematically, erosion operator can be described as [6]

$$\varepsilon\left[f_B(x)\right] = \min_{b\in B} f(x+b) \tag{1}$$

where ε represents the erosion operator, B is the SE and b is the size of SE. A simple example of erosion operator is

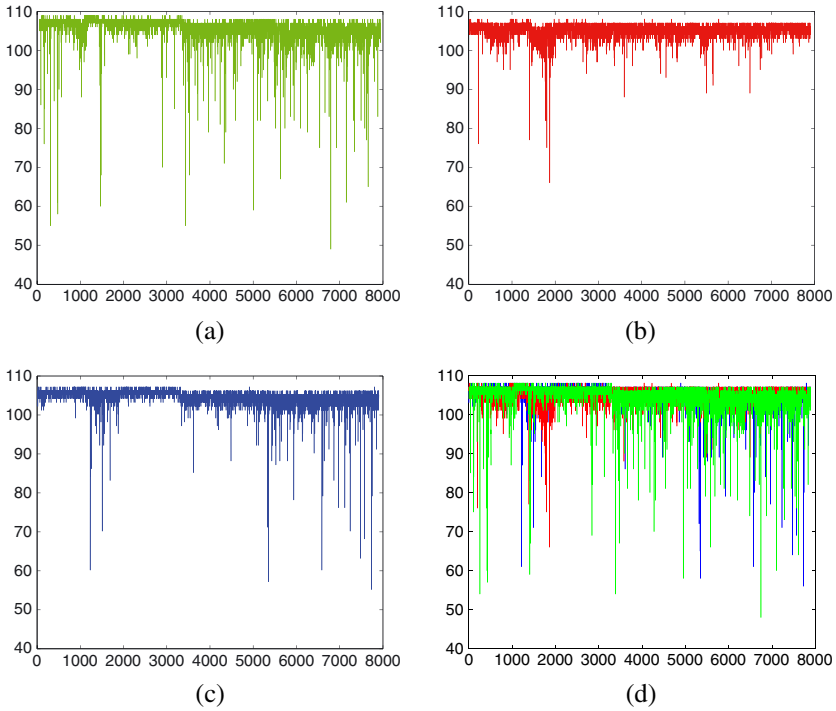

Figure 1. The original link quality indicator data curves.

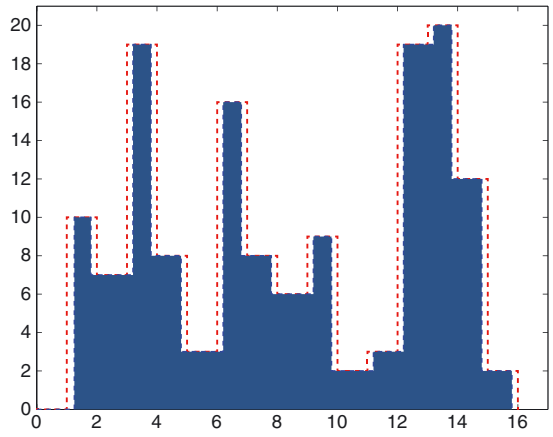

Figure 2. A simple example of erosion operator.

shown in Figure 2. In Figure 2, the dot line is the original data curve, and the shadowed part is the area after being shrunk (eroded).

In this paper, GSD method is applied primarily based on the erosion operator in MM. We use the GSD method to transform the indistinguishable original LQI data into differentiable GSD curve with each curve representing a scenario when different numbers of nodes are active, hence, determine the number of active devices in an IoT system.

3. RELATED WORK

In recent research, data encryption and authentication are the primary two mechanisms to prevent the data from being duplicated and stolen in WSNs [7]. Encryption mechanism can guarantee that the data are incapable to be decrypted by people even if these data have been stolen, while authentication mechanism can guarantee if the data are transmitted from the legitimate nodes and if the data are maliciously modified during transmission. Many protocols based on data encryption and authentication mechanism have been presented in the previous research, such as SPIN protocol [8], TinySec protocol [9] and LiSP protocol [10]. Also, Atakli *et al.* [11] proposed a Weight Trust Evaluation method to monitor and isolate the sensor nodes, which are maliciously interpolated for the purpose of maintaining the security level. Trust-aware recommender system [12] suggests the worthwhile information to the users on the basis of trust.

However, most of them hardly sustain the Intrusion Detection [13,14]. They are incapable of detecting whether "spy nodes" have been placed in a WSN topology. They are unaware of which node and when the WSN topology has been intruded, which constrains the security of the entire WSN. In this paper, an MM-based algorithm is proposed to address the Intrusion Detection problem. We apply this method to the LQI data of each node collected from different sniffers (receivers), transforming the original LQI waveform to the GSD waveform and successfully distinguish the LQI when different number of nodes are communicating in the WSN topology. This paper is an improvement on the basis of our previous research [15], which is the first time that GSD method is applied to the intrusion detection technique.

4. PROPOSED INTRUSION DETECTION METHOD BASED ON GRANULOMETRIC SIZE DISTRIBUTION

In this section, detailed explanation of GSD function is presented. A GSD function is a stepwise function [16]. The establishment of a GSD function is mainly based on the analysis of morphological granulometries.

Granulometries are generated by successively eroding a target curve by enlarging the size of the SEs. Commonly, morphological granulometry performs like sieves [6]. Granulometries are sieved step by step when the size of the SE increases at each step. Hence, the size of granulometries smaller than the SE size will be filtered at each step.

In GSD, the good structure of the image is successively filtered as the size of the mesh (increasing the multiples of SE) increases [6]. The GSD function is generated by continuously using the increasing multiples of SEs to probe (remove) the morphological granulometries in the subgraph until all the subgraph areas have been diminished. A subgraph area of a function $f(x)$ is the area between the X-axis and the curve $f(x)$. Because in our experiment, the original function is translated into the step-function, to be more appropriate and convenient, we regarded a unit square as SE, and the increasing speed of the SE is set to be 1 at each time. Every time the SE cuts the signal waveform into the grains and diminishes part of the subgraph until the entire subgraph area has been removed. A granulometry $\psi(\alpha)$, which is also called the trimming area [17], eroded by the SE Δ with size α is denoted as [17]:

$$\psi(a) = F_{\alpha\Delta}(x) \tag{2}$$

where $F(x)$ is the original LQI data.

Assume a random LQI data $F(x) = (10, 9, 4, 6, 15, 12, 8, 7, 4, 9, 0)$. Figure 3 shows the stepwise function of these LQI data and the granulometry $\psi(\alpha = 2)$.

The GSD function of an LQI waveform can be defined as

$$\{GSD_{\alpha}(F(x))\} = \frac{\sum_{i=1}^{\alpha} \psi(i)}{F_0} \tag{3}$$

where F_0 is the total area of subgraph $F(x)$ and is a positive integer, which will increase by 1 at each time when the subgraph of $F(x)$ is eroded step by step.

The detail procedures of the GSD function are in the succeeding text:

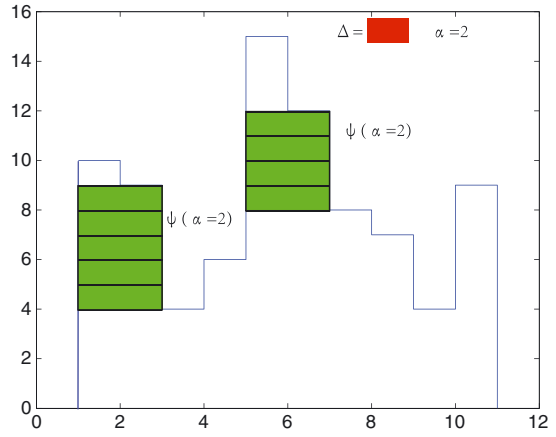

Figure 3. The stepwise function of $F(x)$ and the granulometry when $\alpha = 2$.

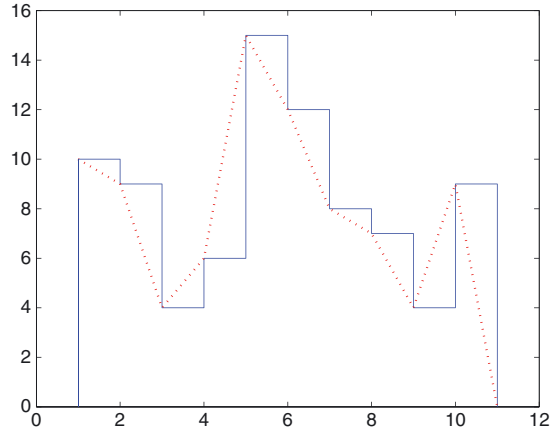

Figure 4. Translating the original signal curve into step-function.

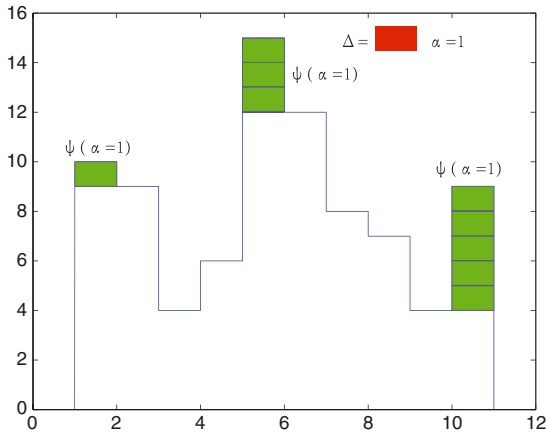

Figure 5. Step (b): Using structuring element to erode every peak of the step-function.

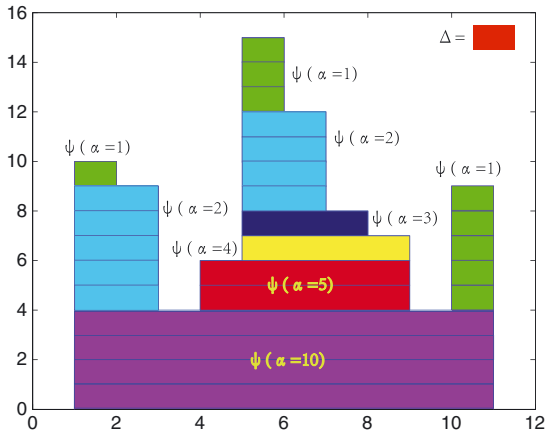

Figure 6. Repeat Step (c) to generate the corresponding matching area.

Step (a): Translate the original signal curve into step-function and calculate the entire subgraph area SG, then goes to Step (b). In this example, the subgraph area SG is 84. Figure 4 shows the original curve (the dot line) and its stepwise curve (the solid line).

Step (b): Use the SE (in our experiment, the SE is a unit square) to erode every peak of the translated step-function. If these peaks match the size of the SE, then count the number of these matching grids and calculate the total area of these matching areas SG_1. Then, remove these matching areas. If no peak matches the SE, regard the eroding area of SE as 0, then goes to Step (c).

Figure 5 shows this step. The shadowed grids are the areas that match the SE. It is easy to count that there are nine unit squares. Hence, the matching area SG_1 is 9.

Step (c): Enlarge the SE by adding one unit square horizontally, which forms a new SE. Using this new SE to erode every peak of the remaining subgraph area. Count the number of these matching grids and calculate these areas. If there is no peak matching the SE, regard the eroding area of SE as 0, then repeat Step (c).

Figure 6 shows the eroding area distribution for different size of the SE after Step (b). By enlarging the SE horizontally by 1 at each time, the corresponding eroding area is generated.

Step (d): Repeat Step (c) until all the subgraph areas have been diminished. Then, we can obtain the eroding areas sequence eroded at each time. Finally, the GSD function can be achieved by calculating the

Table I. Granulometric size distribution (GSD) table for this example.

SE Size	1	2	3	4	5	6–9	10
Number of SE	9	9	1	1	2	0	4
Eroding Areas	9	18	3	4	10	0	40
Cumulative Distribution	9	27	30	34	44	44	84
GSD	9/84	27/84	30/84	34/84	44/84	44/84	84/84

SE, structuring element.

Cumulative Distribution Function (CDF) of this eroding areas sequence. Table I shows the GSD Table for this example, and Figure 7 shows the final GSD function of this example.

In this paper, we apply the GSD method to the LQI data, translating the original overlapped LQI data curve into several distinguishable GSD curves. By directly observing these GSD curves, we know exactly the number of active nodes in a WSN topology. Therefore, we can determine the number of "intruders" in the WSN.

5. EXPERIMENTAL DETAILS

In our experiment, seven TelosB (TPR2400) modules (sensor nodes) were regarded as IoT devices. Four of them were used as sensors, which transmit data normally. Another three were regarded only as receivers in order to record the LQI data when other sensors were communicating. IEEE 802.15.4 standard was used in this experiment; the data rate was 250 kbps, and the size of the packets transmitted was 75 bytes.

The initial experiment consisted of two sensors (N1 and N2) and three receivers. These two sensors were transmitting messages back and forth. Each sensor was monitored, and the LQI value of each received packet was recorded by the three receivers. A base station, which is a TelosB module as well attached to the laptop, assembled the data from all three receivers. Then, an additional sensor (N3) was included, with the sensors now communicating in a ring. The corresponding LQI data were collected by the receivers again. Finally, the fourth sensor (N4) was included, and similarly, the LQI data were recorded for this configuration. In each scenario, we record 20 000 LQI data for each node. Then, we chose the most fluctuating 4000 continuous data to apply with our GSD method. We divided these 4000 data into 10 segments. Each segment has the same number of data. For each segment, we applied

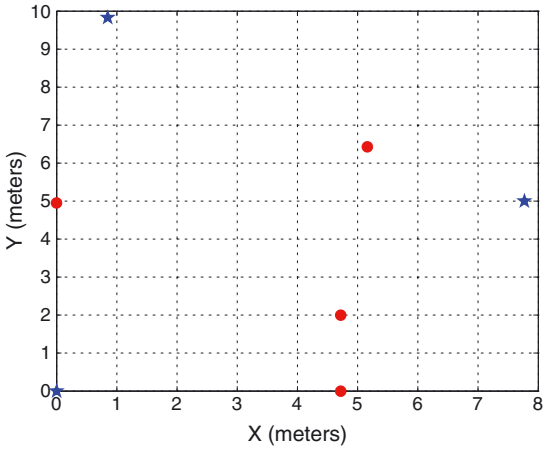

Figure 8. The experiment topology.

our GSD method and generated its GSD curve, hoping to obtain a GSD cluster for each node, in which all the GSD curves follow the same trend while with only a few dissimilarities. The experiment topology is shown in Figure 8. The three star nodes are receivers, and the four circle nodes are sensors.

6. RESULTS ANALYSIS

The primary goal of analyzing GSD function is to monitor the number of active nodes in a WSN, judging whether there are "intruders" in the network. In our experiment, by comparing the GSD clusters of LQI data detected by each sniffer when two nodes, three nodes and four nodes are communicating, it is successful and very clear to differentiate their GSD functions, which make it possible to detect the intruders.

According to the shape-driven characteristic of GSD function, we choose the 4000 successive data points where data fluctuate most frequently, and the shapes of them are

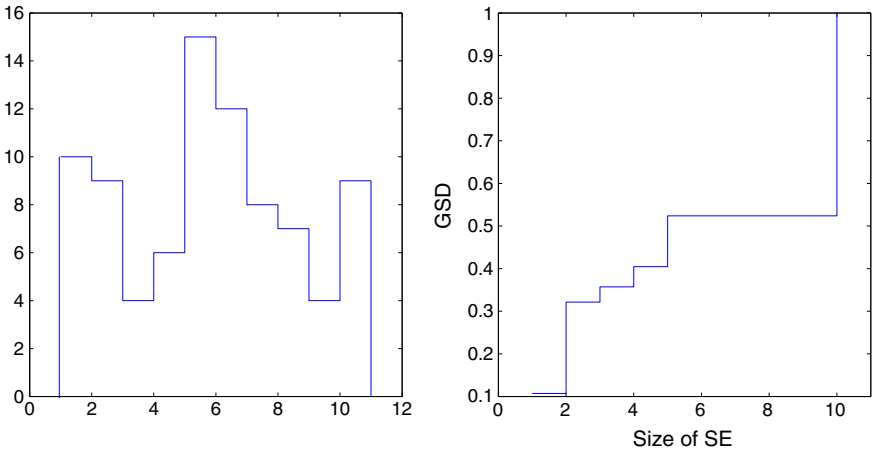

Figure 7. The original signal curve (left) and its granulometric size distribution curve (right).

DOI: 10.1002/sec

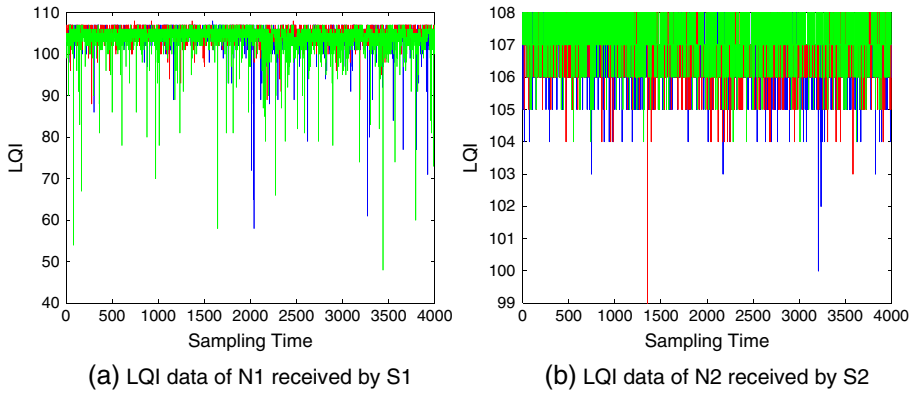

Figure 9. The original overlapped link quality indicator (LQI) data when different numbers of nodes are active. (a) LQI data of N1 received by S1, (b) LQI data of N2 received by S2.

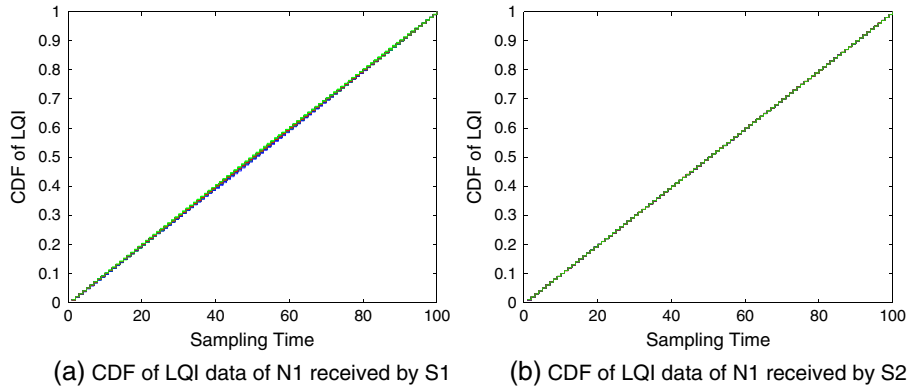

Figure 10. The cumulative distributed function (CDF) of link quality indicator (LQI) data when different number of nodes are active. (a) CDF of LQI data of N1 received by S1, (b) CDF of LQI data of N1 received by S2.

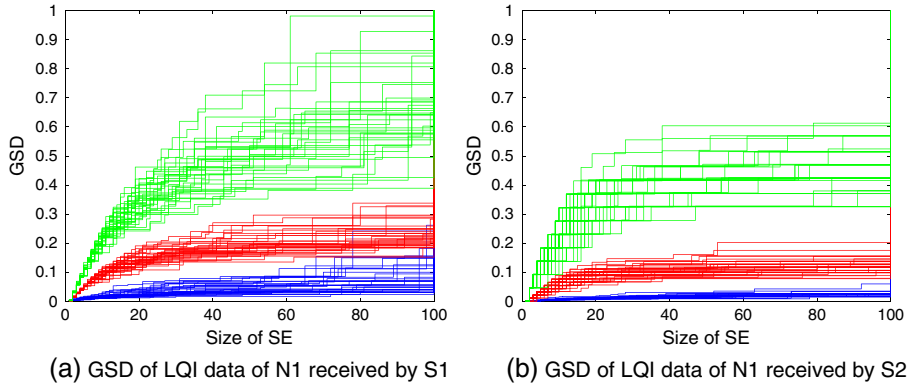

Figure 11. The GSD of link qualtiy indicator (LQI) data when different numbers of nodes are active. (a) Granulometric size distribution (GSD) of LQI data of N1 received by S1, (b) GSD of LQI data of N1 received by S2. SE, structuring element.

dissimilar with each other in order to make sure the final GSD functions can be clearly differentiated. These 4000 data points are divided into 40 segments, which mean we might obtain 40 GSDs for each node when different number of nodes are active. Because the GSD function is a CDF according to its definition, we have compared our GSDs to the CDF of the original data. In Figures 9–11, we have gathered the GSD function of LQI of N1 received by sniffer S1 and LQI of N2 received by S2 when different numbers of sensor nodes are communicating. We have also

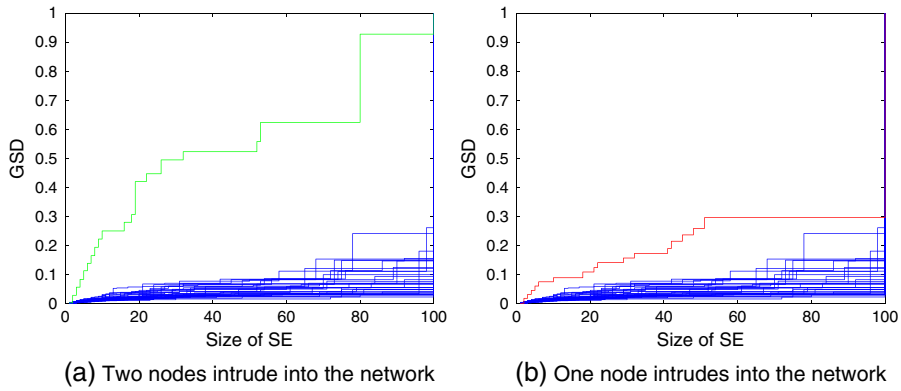

Figure 12. The granulometric size distribution (GSD) of link quality indicator data when the wireless sensor network is intruded. (a) Two nodes intrude into the network, (b) One node intrudes into the network. SE, structuring element.

compared their original LQI data and the CDF of original LQI data together without being applied with GSD.

In Figure 9, the original LQI data curves are overlapped that it is nearly impossible to distinguish them in different scenarios. Moreover, it is very hard to observe their dissimilarities when translating them into CDF. However, by applying our GSD method, it is very clear to observe that the GSD functions have been successfully differentiated into three different clusters. Inside each cluster, the shapes of GSD functions are indeed quite similar, while an apparent gap is generated between each two clusters when different number of nodes are communicating, which makes it possible to determine the number of nodes by briefly observing and analyzing the shape of GSD functions. Although the environment of the WSNs topology may change over time, the LQI of each node will change as well, leading to the change of final GSD functions. In a stable or a hush environment, the environmental effects on each node can be simply neglected by applying our GSD method if each node is placed in the same environment.

Assume a wireless topology consists of three wireless sensor nodes. At a certain moment, an unknown malicious node suddenly intrudes into this topology. In Figure 12, by directly observing the updated LQI original data or their CDF, it is impossible to know whether there is a malicious node intruding into the network. However, based on the experiment results mentioned earlier, by translating the original LQI data into the GSD function, it is possible to notice this malicious node. In Figure 12, at the beginning of the simulation, the number of nodes turned on to transmit data is known as 2. When nodes intrude in, the GSD curve of LQI will suddenly change. By observing this updated GSD, the number of intrusive nodes can be determined. Throughout the entire GSD generation procedures, the original data are intact, which make our GSD method very appropriate for the intrusion detection.

7. CONCLUSION

In this paper, a practical problem in IoT system is presented, which is how to monitor the network to prevent it from being intruded. GSD-based method is proposed to solve this problem. The key point of GSD is to translate the original indistinguishable LQI data into several differentiable curves without any change of original data. The mathematical expression of GSD and procedures to generate GSD function are presented in detail in this paper. Finally, by comparing our GSD method to the CDF of the original LQI data, our experiment results indicate that it is possible to determine the number of active nodes clearly in an IoT system in real-time without any change or loss of the original data, hence, implementing the intrusion detection. In future work, how to localize the intrusive nodes might be considered after the number of them has been determined. In addition, how to make a global criterion between each GSD cluster and how to automatically make efficient observation by computers without people, especially when the number of nodes becomes large, is another task in our future research.

ACKNOWLEDGEMENTS

This work is supported by the 2014 Guangzhou Pearl River New-star Plan of Science and Technology Project (No. 2014J2200023), the Open Fund of Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis (No. GDUPTKLAB201304), the 2014 Shenzhen City “Knowledge Innovation Program" Project (No. JCYJ20140417113430604) and the Open Project Foundation of Information Technology Research Base of Civil Aviation Administration of China (NO. CAAC-ITRB-201406).

REFERENCES

1. Ming-Whei F. Wireless sensor network industrial view? What will be the killer apps for wireless sensor network? *IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing, SUTC’ 08*, Taichung, Taiwan, 2008; 270.

2. Nouha B, Anis K, Luca M, Marco AZ, Habib Y, Carlo AB, Mário A. Radio link quality estimation in wireless sensor networks: a survey. *ACM Transactions on Sensor Networks* 2012; **8**(4): 33, Article 34.
3. Tao L, Alberto EC. Data-driven link quality prediction using link features. *ACM Transactions on Sensors Networks* 2014; **10**(2): 35, Article 37.
4. Guoan H. A link quality evaluation model based on the three-dimensional space in wireless sensor network. *Information Technology Journal* 2014; **13**: 720–724.
5. Matheron G, Jean S. The birth of mathematical morphology. *In Proceedings 6th International Symposium Mathematical Morphology*, Sydney, Australia, 2002; 1–16.
6. Guardiola IG, Mallor F. A nonparametric method for detecting unintended electromagnetic emissions. *IEEE Transactions on Electromagnetic Compatibility* 2013; **55**(1): 58–65.
7. Simplicio, Jr, MA, De Oliveira BT, Margi CB, Barreto PSLM, Carvalho TCMB, Mats N. Survey and comparison of message authentication solutions on wireless sensor networks. *Ad Hoc Networks* 2013; **11** (3): 1221–1236.
8. Adrian P, Robert S, Tygar JD, Victor W, David EC. SPINS: security protocols for sensor networks. *Wireless Networks* 2002; **8**(5): 521–534.
9. AlMheiri SM, AlQamzi HS. Data link layer security protocols in wireless sensor networks: a survey. *10th IEEE International Conference on Networking, Sensing and Control (ICNSC)*, Paris-Evry, France, 2013; 312–317.
10. Taejoon P, Kang GS. LiSP: a lightweight security protocol for wireless sensor networks. *ACM Transactions on Embedded Computing Systems* 2004; **3** (3): 634–660.
11. Idris MA, Hongbing H, Yu C, Wei SK, Zhou S. Malicious node detection in wireless sensor networks using weighted trust evaluation. In *Proceedings of the 2008 Spring Simulation Multiconference (SpringSim '08)*. Society for Computer Simulation International: San Diego, CA, USA, 2008; 836–843.
12. Yuan W, Shu L, Chao HC, Guan D, Lee YK, Lee S. ITARS: trust-aware recommender system using implicit trust networks. *IET Communications* 2010; **4**(14): 1709–1721.
13. Moorthy M, Sathiyabama S. A study of intrusion detection using data mining. *2012 International Conference on Advances in Engineering, Science and Management (ICAESM)*, Nagapattinam, India, 2012; 8–15.
14. Zhang YY, Chao HC, Chen M, Shu L, Park CH, Park MS. Outlier detection and countermeasure for hierarchical wireless sensor networks. *IET Information Security* 2012; **4**(4): 361–373.
15. Yanwen W, Xiaoling W, Hainan C, Guangcong L, Lei S. A new intrusion detection method for malicious attack in WSN based on mathematical morphology. *9th International Conference on Communications and Networking in China, Chinacom 2014*, Maoming, China, 2014.
16. Qing-hua Z, Yu-xian Z, Ping Y, Ji-jun X. The Analysis about the unit-step function of the bending moment equation of girder. *Intelligent Computing and Integrated Systems (ICISS), 2010 International Conference on*, Guilin, China, October 22–24, 2010; 479–481.
17. Yanwen W, Guardiola IG, Xiaoling W. RSSI and LQI data clustering techniques to determine the number of nodes in wireless sensor networks. *International Journal of Distributed Sensor Networks* 2014; **2014**: 11, Article ID 380526.