Building Trusted Routing with Subjective Logic in Wireless Sensor Networks

Zhou Hongwei[1, a], Yuan Jinhui[1, *b]

[1] Information Engineering University, Zhenzhou, 450004, China

[a]email: hong_wei_zhou@hotmail.com, [b]email: jcyjh@126.com, corresponding author

Keywords: wireless sensor network; routing; subjective logic; spatial correlation

Abstract. Since sensor nodes are usually deployed in a hostile environment, it is easy to be physically captured by an adversary. Thus, an adversary is capable of disturbing the entire network. To overcome it, we propose a trusted routing algorithm using subjective logic. In our solution, every node checks the similarity between itself and its neighbor nodes which are denoted as subjective opinion. On the opinion, the node choose its aggregation node in its neighbor nodes. Furthermore, if one aggregation node is compromised by an adversary, its neighbor nodes report the alarm accompanying with a subjective opinion and the routing is adjusted on the fusion reported opinions. Thus, the compromised nodes are circled. Our discussion and partly implementation show that our scheme provides trusted routing and improves the accuracy of sensing data without heavy energy overhead.

Introduction

Sensor nodes are not usually physically protected, and they are easy to be compromised by an adversary. With the compromised nodes, an adversary are capable of inserting the false data to subvert the sensor network. For example, wireless sensor network is deployed around the house, and the user are aware of changing stresses with the support of these nodes. However, if an adversary capture several nodes, and send false data to the user, the user fail to obtain the precise sensing data. Furthermore, to save the energy, data aggregating technology are used widely. So it is possible that an adversary tampers a collection of sensing data only capturing one node. Due to the limited cost, it is difficult to develop complex security mechanism for wireless sensor network to defeat this attack.

There are some work to defeat the node compromised attack. Reputation-based framework are presented which evaluate every node's trust on their activity[1,2]. Our previous work[3] introduces subjective logic to check the false data sent by compromised node. However, the solutions do not eliminate the false data in wireless sensor network, and does not adjust the routing. In our opinion, due to the limited energy, the false data should not be transmitted because message transmission consume the main energy of the nodes. Moreover inserted false data disturb the user and reduce the accuracy of sensing data. In our opinion, the routing should circle the compromised node to avoid the interference of false data.

To overcome the above limitation, in this paper, we present an trusted routing algorithm that is design to expose and circle compromised node, and rebuild the routing without heavy overhead. How to identify the false data? We identify them by cooperation of neighbor nodes with the support of the subjective logic. In other words, every neighbor node poses the opinion that the sensing data is a false data, and the conclusion is drawn on the fused opinion. To circle the compromised node, these neighbor nodes cooperate to choose a new aggregation node. Our discussion and partly implementation show that our scheme provides trusted routing and improves the accuracy of sensing data without heavy energy overhead.

The rest of the paper is organized as follows. Section 2 presents the backgrounds of our solution, and section 3 describes our solution. Section 4 presents the detail of our discussion. At last, we concludes this paper.

Background

In wireless sensor network, nearby nodes usually report similar data, and it is usually called spatial correlation. This is a solid ground for data aggregation and energy saving. For example, there are ten sensor nodes deployed in the building to obtain the temperature. Normally, their sensing data are similar. So it is not necessary that every node report its data every time. Thus, to save the energy, these nodes may construct a cluster, and only one report data to sink node. On spatial correlation, as a example, [5] develops a clustering-based aggregation algorithm that is still capable of providing accurate data. Spatial correlation can be also used to expose false data. On the above example, if nine node's data are similar, but one's data are sharply different from others, we can draw a conclusion that the node may be compromised and report a false data. Some work detect false data with spatial correlation[3,6]. Our solution also detect compromised node with the same features.

In this paper, we use subjective logic to find compromised node. Subjective logic, which is a kind of probabilistic logic, is proposed by audun jøsang[7]. Its main feature is taking uncertainty into account, and it is suitable to model the situation without fully knowledge. For example, it can be used to analysis trust network[8]. Subjective opinion is a main concept of subjective logic. A binomial opinion about the proposition x is denoted as Wx={b,d,u,a} where b is belief that proposition x is true, and d is belief that proposition x is false, and u is the amount of uncertain, and a is the priori probability. For example, a proposition y is describe as follows: tomorrow may be fine. A forecaster hold an opinion Wy={0.5,0.3,0.2,0.5}. On the opinion, we know the forecaster believes that the chance of y is 60% with a formula $exp=b+a*u$. Fusion operator is a main operator of subjective logic, and we use it to fuse these opinions on the same proposition. For example, W^A_x={0.5,0.3,0.2,0.5} means user A hold an opinion on proposition x, W^B_x={0.6,0.1,0.3,0.5} is another opinion hold by user B on the same proposition. Fusing W^A_x and W^B_x, we can produce a fusion opinion W^{AB}_x={0.61,0.25,0.14,0.5} according to the fusion rule of subjective logic. In this paper, we use fusion operator to fuse the neighbor's opinions, and present the expectation value of the fusion opinion. The expectation value is used to detect compromised node.

Our solution

In this paper, we suppose the sensor network include n sensor nodes and one sink node for simple discussion. However, we believe that it is easy to be extended to other situations. Our solution comprise of several parts: exploring location, selecting aggregation node, completing routing and circling compromised node. We discuss them as follows.

Exploring location

Every node has to locate the distance to sink node before building the routing. To this end, the sink node flood the location message m_0. These nodes, which is deployed near the sink node, are received the message m_0. We denoted these nodes as 1-nodes. Like sink node, 1-nodes flood the message m_1, and some other nodes can be capture the message m_1. Noticed that some 1-nodes also observe the message m_1, but they discard the messages. These nodes that receive message m_1 are denoted as 2-nodes. Similar to this, all nodes can be denoted as i-node.

It is necessary to locate any node in sensor network for build the routing. An k-node has to find a road k-1-node→k-2-node→... →1-node to reach sink node. In some time, an k-node stop to work, for example it exhausts its energy, some k+1-nodes may find other node as their bridge to sink node. If these nodes fail to find any available k-node, it means that sink node receives no more than their sensing data, and we fail to build the routing for the network. On the other side, it is possible that the radio link may break. However, in the paper, we suppose that the links are reliable except that the node exhausts its energy. In this step, all nodes flood one location message.

Selecting aggregation node

In this paper, we use subjective opinion to score the spatial correlation between two nodes. As mentioned earlier, subjective opinion is a 4-tuple {b,d,u,a}. So the key point is to build the map from spatial correlation to {b,d,u,a}.

Suppose node *A* and node *B* are deployed, and we observe and remember their sensing data *k* times. Their data are respectively $\{a_1,a_2,...,a_k\}$ and $\{b_1,b_2,...,b_k\}$. With these data, we can obtain the average of the difference of their data as follows which is denoted as β.

$$\beta = \frac{\sum_{i=1}^{k} |a_i - b_i|}{k}$$

On β and the data of k+1 times, we present the map as follows.

case 1: if $|a_{i+1} - b_{i+1}| \le \beta$

$$\begin{cases} b = \dfrac{\beta - |a_{i+1} - b_{i+1}|}{\beta} \\ d = 0 \\ u = \dfrac{|a_{i+1} - b_{i+1}|}{\beta} \\ a = 0.5 \end{cases}$$

case 2: if $2\beta > |a_{i+1} - b_{i+1}| > \beta$

$$\begin{cases} b = 0 \\ d = \dfrac{|a_{i+1} - b_{i+1}| - \beta}{\beta} \\ u = \dfrac{2\beta - |a_{i+1} - b_{i+1}|}{\beta} \\ a = 0.5 \end{cases}$$

case 3: if $|a_{i+1} - b_{i+1}| > 2\beta$

$$\begin{cases} b = 0 \\ d = 1 \\ u = 0 \\ a = 0.5 \end{cases}$$

For example β=1, a_{k+1}=20.3, b_{k+1}=21.1, we obtain subjective opinion W={0.2,0,0.8,0.5}. Furthermore, we can obtain the expectation value of the above opinion is 0.5. There are a preset parameter δ. If opinion's expectation value is bigger than δ, we consider these nodes are spatial correlated, and one node may be aggregation node of other node. δ is so important because the capability of detecting compromised node depends on it.

To simple the routing algorithm, every node has a unique ID. With a unique ID, it is easy to choose the aggregation node. In this paper, the node with a bigger ID is selected as the default aggregation node. However, there is an excepted situation that the default node has insufficient energy to act as an aggregation node.

Noted that there are three situations to select one new aggregation node. First, a node finds its aggregation node is no more than its spatial correlation node, it has to select other node as its aggregation node, or it work as a new aggregate node. Second, with the cooperation of its neighbor nodes, a node consider its aggregation node is compromised, and it want to circle the compromised node. Last, an aggregation node almost exhaust its energy, and it is not suitable to act as aggregation node. However, the method to select the new aggregation node is same.

On the above discussion, we can present our algorithm to select the aggregation node.

Algorithm 1: selecting aggregation node algorithm

(1) Every node flood the messages{ID, sensing data, energy level} k times.
(2) At k+1 time, every node try to find its spatial correlated node with above introduced method on subjective logic, which has enough energy to act as aggregation node.
(3) If a node find a spatial correlated neighbor node who has a bigger ID, it stops sensing data and log its spatial correlated node's ID.
(4)If a node does not find a spatial correlated neighbor node or its spatial correlated neighbor node has a less ID, the node act as an aggregation node.

Completing routing

After selecting aggregation node, every aggregation node has to find a road to reach sink node. To this end, every aggregation node floods a completing routing message {ID, *i*-node} where *i*-node is marked in exploring location phase. The *i*-1-node, who observe the message, flood the asking message {ID, *i*-1-node, Y/N} where Y/N means whether the node is one aggregation node. There is a situation that several nodes all ask the completing routing message. To save energy of the node, aggregation node should act as routing node to complete the routing. When more than one aggregation node ask the message, we select the node with a bigger ID as a completing routing node. If no any node is not the aggregation node, the node with a bigger ID is considered as a completing

routing node. Noted that a node does not ask any message when it has little energy to save energy.

We present completing routing algorithm as follows.

Algorithm 2: completing routing algorithm

(1) Every aggregation node A flood messages{ID, i-node}.
(2) All i-1-node who observe above message send message{ID, i-1-node, Y/N}.
(3) If a i-1-node is aggregation node, A consider it as a completing routing node. If more than 2 i-1-nodes are aggregation node, A consider the node with a bigger ID as a completing routing node.
(4)If there is not a i-1-node is aggregation node, A consider the node with a bigger ID as a completing routing node.
(5)If there is not any node ask the request message, we fail to complete routing.

Circling compromised node

It is possible that an adversary suddenly physically capture one node and insert a false data into the network. In our routing algorithm, every node are always to observer the sensing data of its neighbor node, and check the integrity of its data. If a node find its aggregation node sends a false data, it will flood an alarm message including an alarm and its data. Next, with the cooperation of other neighbor nodes, the compromised node is detected and the alarm message is flooded to more related nodes.

We illustrate our idea with an example in fig 1. The current route road is node 5→node 4→node 1→sink node. At some time, node 4 is captured by an adversary, and send a false data. So its neighbor nodes, including node 1, node 2, node 3 and node 5, detect node 4 may send a false data. So these neighbor nodes flood an alarm message{compromised node' ID, subjective opinion}. Suppose node 2 can receive the messages sent by node 1 and node 5, but node 3. Thus node 2 fuses three opinions to judge whether node 4 send a false data. If node 2 consider that node 4 send a false data, node 2 floods selecting aggregation node message to find a new aggregation node.

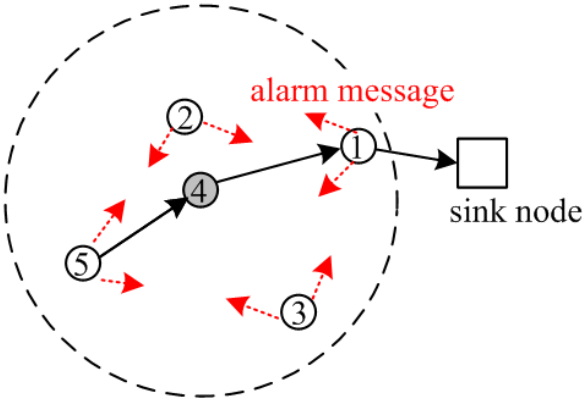

Fig.1. The example of detecting the compromised node

The remaining problem is how to fuse the opinions of neighbor nodes. To the end, we fuse the different opinions with the fusion operator of subjective logic. The detail of fusion operator is described as follows.

Suppose ω_x^A and ω_x^B are the opinions hold respectively by subject A and subject B on the same proposition x where $\omega_x^A = \{b_x^A, d_x^A, u_x^A, a_x^A\}$ and $\omega_x^B = \{b_x^B, d_x^B, u_x^B, a_x^B\}$, $\omega_x^{A,B}$ is the fusion opinion, and $\omega_x^{A,B} = \omega_x^A \oplus \omega_x^B$. Suppose that $\omega_x^{A,B} = \{b_x^{A,B}, d_x^{A,B}, u_x^{A,B}, a_x^{A,B}\}$, $k = u_x^B + u_x^A - u_x^A u_x^B$, the fusion rule is summary:

case 1: $k \neq 0$ case 2: k=0

$$\begin{cases} b_x^{A,B} = \dfrac{(b_x^A u_x^B + b_x^B u_x^A)}{k} \\ d_x^{A,B} = \dfrac{(d_x^A u_x^B + d_x^B u_x^A)}{k} \\ u_x^{A,B} = \dfrac{u_x^A u_x^B}{k} \\ a_x^{A,B} = \dfrac{a_x^A u_x^B + a_x^B u_x^A - (a_x^A + a_x^B) u_x^A u_x^B}{k - u_x^A u_x^B} \end{cases} \qquad \begin{cases} b_x^{A,B} = \dfrac{b_x^B + b_x^A \gamma}{\gamma + 1} \\ d_x^{A,B} = \dfrac{d_x^B + d_x^A \gamma}{\gamma + 1} \\ u_x^{A,B} = 0 \\ a_x^{A,B} = \dfrac{\gamma a_x^A + a_x^B}{\gamma + 1} \end{cases}, \quad \gamma = lim\left(\frac{u_x^B}{u_x^A}\right)$$

We also illustrate the process with the example of fig 1. Node 1, node 4, and node 5 are the neighbor nodes of node 2, and node 2 is capable of receiving their flood message. For node 2, there are three opinions to be fused including its opinion, and we denote these opinion as W^1_x , W^2_x and W^5_x. The fusion opinion is computed as follows: $W = W^1_x \oplus W^2_x \oplus W^5_x$. At last, node 2 obtain the expectation value of *W*. If the value is bigger than a preset value Ø, node 2 considers that node 4 may be a compromised node. After detecting compromised node, the nodes who lose their aggregation node have to select a new aggregation node. The process is same to selecting aggregation node algorithm, and we do not discuss this.

We conclude how to circle compromised node.

Algorithm 3: circling compromised node algorithm
(1) Every node observes its aggregation node's data. (2)If a node finds its aggregation node may be compromised, it flood an alarm message{compromised node's ID, subjective logic}. (3) The node fuses all subjective opinions on the same node to obtain the finally opinion. (4) If the expectation value of fused opinion is bigger than Ø, the node selects a new aggregation node allowing selecting aggregation node algorithm.

Discussion

The first interesting problem is how many messages are necessary to circle the compromised node and rebuild the routing. In our opinion, for every node, it is necessary to send two messages to circle the compromised node. Besides that, in worst case, every node has to send one message to complete the routing. Fortunately, the worst case is hardly appear. In the worst case, all nodes are deployed intensively, and every node answer the request message sent by one node. Usually, it is impossible to deploy sensor network like this. In our opinion, if the compromised node has *n* neighbor node, about 2*n* messages have to be sent for circling the compromised node, and about *m* message to complete the routing which *m* is the number of the node which is the number of neighbor nodes of the new aggregation node which are more close to sink node. We optimistically consider that about 3*n* messages are need to rebuild the routing.

The second interesting problem is the accuracy of the sensing data. We suppose there are 100 nodes are deployed in the situation, and only 9 node work as the aggregation nodes. Furthermore, these nodes had sent data to sink node 100 times. Now, if an compromised aggregation node always send false data to sink node as the representative of 7 nodes at 30 times. We can estimate the deviation of sensing data. For example, in ideal situation, every node sensing data is same, and false data is increased by 10%. The sum of real sensing data is α, while the sum of all sensing data, including false data and other real data, is 1.05*α. Our algorithm can detect the false data, and remove it from the network for improving the accuracy of sensing data. In ideal situation, the false data is only transmitted one time.

We discuss further on Intel Berkeley research lab data set[9]. In this situation, there are 54 nodes deployed in the lab. In this paper, we only discuss the temperature which is one of four sensing kind of data. And the similar conclusion can be drawn on other kind data of the same data set. First, we

want to know how many neighbor nodes for every node in the network. We set the most distance between the node and its neighbor node as the 4, and the sum of neighbor nodes of every node is 46. When the distance is set to 10, the sum is 438. This means that every node has about 5 neighbor nodes, and it takes about 15 messages to rebuild the routing for circling the compromised node. On the other side, when one node is compromised by an adversary, its neighbor nodes are capable of detecting it with our partly simulation implementation on the above data set. For example, node 51 has 4 neighbor nodes when the distance is set 7. When the false data is forged by increase the normal data by 20%, it is evident that neighbor nodes expose the compromised data. However, it is possible to miss some false data if an adversary forge carefully it. To overcome it, we have to adjust the parameter of the algorithm. However, we consider it is out of this paper.

Conclusion

In this paper, we present a trusted routing algorithm using subjective logic. Our solution comprises of four sub-algorithms including exploring location, selecting aggregation node, completing routing and circling compromised node. Our discussion and partly implementation show that our algorithm take about 3*n* messages to rebuild the routing which *n* is the number of neighbor nodes of compromised node, and it is capable of exposing compromised node to improve the accuracy of sensing data.

Acknowledgement

The authors would like to thank the anonymous reviewers for their insightful comments that helped improve the presentation of this paper. The work is supported in part by the National Natural Science Foundation of China (61303074).

References

[1] S. Ganeriwal and M. Srivastava. Reputation-based framework for high integrity sensor networks. In ACM SASN, October 2004.

[2] Y. Sun, Z. Han, W. Yu, and K. Liu. A trust evaluation framework in distributed networks: vulnerability analysis and defense against attacks. In IEEE INFOCOM, April 2006.

[3] Jinhui Yuan, Hongwei Zhou, and Hong Chen, Subjective Logic-Based Anomaly Detection Framework in Wireless Sensor Networks, International Journal of Distributed Sensor Networks, vol. 2012,

[4] Luo H, Luo J, Liu Y, et al. Adaptive Data Fusion for Energy Efficient Routing in Wireless Sensor Networks[J]. IEEE Transactions on Computers, 2006, 55(10):1286-1299.

[5] Ma Y, Guo Y, Tian X, et al. Distributed Clustering-Based Aggregation Algorithm for Spatial Correlated Sensor Networks[J]. IEEE Sensors Journal, 2011, 11(3):641-648.

[6] Chatzigiannakis V, Papavassiliou S. Diagnosing Anomalies and Identifying Faulty Nodes in Sensor Networks[J]. IEEE Sensors Journal, 2007, 7(5):637-645.

[7] AUDUN JØSANG. A logic for uncertain probabilities. International Journal of Uncertainty Fuzziness and Knowledge-Based Systems, 2001,9(3), 279-311.

[8] Jsang A, Bhuiyan T. Optimal Trust Network Analysis with Subjective Logic. Second International Conference on Emerging Security Information, Systems and Technologies. IEEE Computer Society, 2008:179 - 184.

[9] Intel Berkeley Research Lab, http://berkeley.intel-research.net/labdata/.