An Extended PEGASIS Protocol Based on Group Authentication with Merkle Tree in WSNs

Wangsheng Fang¹, Xiangyu Wang¹, Haomin Xu¹

¹ School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou, China

Abstract. With the constant improvement of Wireless Sensor Network, its application scenarios become more diverse. Due to the limited resources of sensor nodes, there are many researches devoted to the optimization of network structure and the prolongation of network life. The proposed extended protocol is based on the chain-based routing technique of PEGASIS. In a newly deployed network, intra-group trust is formed through the Group Authentication based on Merkle Tree, which can protect the node information and resist the deception of false message. In the steady phase, the whole network re-chain is replaced by re-chain within the group when any node dies. Simulation results show that the scheme proposed in this paper has obvious improvement in terms of residual energy and network life compared with PEGASIS.

1. Introduction

CWSN (Chain-type Wireless Sensor Network) is a kind of WSNs (Wireless Sensor Network) that a number of sensor nodes are deployed in a long and narrow zone using a chain-type network topology [1]. As shown in figure 1, it’s an example of CWSN. Sensor nodes are constrained by limited energy consumption, insufficient computing power, small storage capacity and finite communication radius [2].

![Figure 1. An example of chain-type wireless sensor network](image-url)

PEGASIS (Power-Efficient GAthering in Sensor Information Systems) is one of the most well-known routing protocols [3]. Some of the most compelling studies have focused on the energy consumption optimization of CWSNs. PEGASIS-INL only selects the node as the leader within the communication range of the base station, and uses multi-overlapping chain formation and data aggregation to transmit the message to the BS (Base Station) [4]. PEGASIS-ACO uses Ant Colony Optimization to select the nearest node as the next-hop node, which greatly reduces the energy consumption of the entire network [5]. However, the works above are all energy-saving deployment in the initial stage of the network. Whenever a node dies, the BS releases control message to start a new round, which will cause unnecessary energy consumption loss undoubtedly. There are a few works aimed at lightweight IoT (Internet of Things) devices, which store transaction information within the
network in the form of Merkle tree to form local trust between nodes. This method not only alleviates the problem of resource limitation of nodes, but also improves the security of nodes [6][7].

Considering the lower initial energy and a smaller communication range of sensor nodes, we propose an extended PEGASIS protocol based on Group Authentication with Merkle Tree (PEGASIS-GAMT). Nodes are grouped on the basis of the optimal value of group, then they calculate the Hash of the Merkle Root using node identity information in the same group and store it. When a node dies in the network, it will no longer release control information through the base station and re-chain, but the members of the group where the node dies will verify each other and re-chain within the group. Compared with the simulation results of energy consumption in literature [3], the PEGASIS-GAMT proposed in this paper has obvious improvement in terms of energy consumption.

2. Scenario Description

2.1. Network model of PEGASIS

As shown in figure 1, many sensor nodes are deployed in a long and narrow area with chain type network topology. The nodes close to the BS will also transmit data as the relay node for the nodes far away from the BS. The arrow in figure 1 represents the data transmission. We define the initial energy of each sensor node as $E_{init}$, and the distance between node $SN_i$ and $SN_{i+1}$ as $d_i$. Particularly, BS is the next-hop of node $SN_n$. $L$ and $H$ denote the length and width of the sensing region of the CWSN.

2.2. Energy Model of Wireless Communication

For our simulation, we assume a simple model for the radio hardware energy dissipation used in literature [8]. As shown in figure 5, the energy consumption generated by the transmitter is mainly composed of signal transmission and power amplifier, while the energy consumption of the receiver is only generated by signal reception.

According to the potential data redundancy, it performs data aggregation to reduce the unnecessary noise. The equations used to calculate transmission and receiving costs for a $l$-bit message and a distance $d$ are shown below:

\[
\begin{align*}
E_{TX}(l, d) &= E_{TX_{elec}}(l) + E_d(l) + E_{TX_{amp}}(l, d) \\
E_{TX}(l, d) &= E_{elec} \ast l + E_d \ast 2l + \varepsilon_{amp} \ast l \ast d^2
\end{align*}
\]

Receiving

\[
\begin{align*}
E_{RX}(l) &= E_{RX_{elec}}(l) \\
E_{RX}(l) &= E_{elec} \ast l
\end{align*}
\]

The electronics energy, $E_{elec}$, depends on the length of message, whereas the amplifier energy, $\varepsilon_{amp}$, depends on the distance to the receiver. We assume the radio channel is symmetric, the energy consumption of unit data transmitted from node $i$ to node $j$ is the same as that of unit data transmitted from node $j$ to node $i$ for a given SNR (Signal to Noise Ratio).

2.3. Merkle Tree

Merkle Tree is a hash binary tree, which is a data structure used for fast generalization and verification of large-scale data integrity. Traditional Merkle Tree is constructed from the bottom up. As shown in figure 3, Merkle tree is constructed from the identity of sensor nodes $N_1$, $N_2$, $N_3$ and $N_4$. Firstly, leaf nodes calculate the hash using the corresponding sensor node identity and store the hash value. Then,
the hash value of the parent node is calculated all the way to the top node, which we call it as Merkle Root.

![Merkle Tree Diagram](image)

**Figure 3. A simple example of Merkle Tree**

Merkle Tree utilizes one-way hash function to design security authentication method, which can hide the real identity information of nodes and prevent attackers from using false data to conduct malicious spoofing [9].

### 3. Our Proposed Strategy

#### 3.1. Optimum Number of Group

In PEGASIS, the BS will broadcast control information to make the whole networker-chain when a node dies. This method only loses part of the perception information of the current rotation, but also causes a large amount of energy waste of nodes. For these considerations, the grouping authentication based on Merkle Tree proposed in this paper establishes local trust intra-group members. When a member node in a group dies, the entire network no longer needs to re-chain, the nodes in that group will re-chain after the verification. Therefore, it is very important to select the optimal value of the number of group members. Assuming that $N$ nodes are uniformly distributed in the perceptual region of $H^*L$, $k$ groups are divided into per round, the equation of energy consumption required by each group to verify each other is below:

$$E_a = 2 \left( \frac{N}{k} - 1 \right) \times l_a \times E_e + \left( \frac{N}{k} - 1 \right) \times l_a \times \varepsilon_{amp} \times d^2$$

$$E_a = 2 \left( \frac{N}{k} - 1 \right) \times l_a \times E_e + \left( \frac{N}{k} - 1 \right) \times l_a \times \varepsilon_{amp} \times d^2$$

Where $d$ is the distance between nodes, and it is assumed that nodes are randomly distributed in the sensing region,

$$E[d^2] = \iint (x^2 + y^2) \rho(x, y) dxdy = \iint r^2 \rho(r, \theta) r dr d\theta$$

If we assume that the sensing region is a circle which radius $R = \sqrt{HL/\pi k}$ and $\rho(r, \theta)$ is constant for $r$ and $\theta$, the density of nodes $\rho = 1/((HL)^2/k)$, (6) simplifies to

$$E[d^2] = \frac{1}{2\pi} \frac{HL}{k}$$

Therefore, the total energy for the frame is

$$E_{total} = kE_{group} = N(2lE_e + 2lE_d) + (N - k)l \varepsilon_{amp} \times d^2$$

$$E_{total} = kE_{group} = N(2lE_e + 2lE_d) + (N - k)l \varepsilon_{amp} \times d^2$$

After the derivative of $E_{total}$ is set with respect to $k$ to zero, the optimum number of group will be found.
\[ E_{\text{total}}' = l \varepsilon_{\text{amp}} d^2 - \left[ \frac{N e_{\text{ampHL}}(l-l_a)}{2\pi} + 2N^2 l_a E_e \right] \frac{1}{k^2} - \frac{2N^2 l_a e_{\text{ampHL}}}{2\pi} \frac{1}{k^3} = 0 \] (9)

According to the root formula of cubic equation with one variable, the optimal value range of real root \( k \) can be obtained, which is varied between 1 and 10. However, in order to make it clearer, we choose the range between 4 and 10 (the curve decreases monotonically between 1 and 4), and the optimal value of grouping number is finally selected according to the polyline graph, as shown in figure 4.

Figure 4. The optimum number of group

In PEGASIS, the number of groups varies between 1 and 10 with the average energy consumption of per round. Figure 4 shows that the optimal range of energy consumption ranges from 5 to 7. When the number of groups is 6, the optimal energy consumption is reached in the CWSN with 30 sensor nodes. When the number of members \( N_{\text{num}} = [N/k] \) in a group is 5, the energy utilization rate is the highest. This is consistent with the efforts to reduce communication energy consumption in various protocols of CWSN. The results also verify the analysis in this paper.

3.2. Node Grouping Scheme

In the initial phase, PEGASIS-GAMT is similar to the traditional PEGASIS, just forming the group through the optimal value calculated above. Extension is only shown when nodes start to die. The general process of node grouping scheme in the initial phase is as follows:

1. Node \( i \) calculates hash value \( H(ID_i) \) and stores it.
2. Node \( i \) broadcasts node identity information, which includes location hash value and initial energy \( \{x_i, y_i, H(ID_i), E_{\text{init}}\} \).
3. Node \( i \) receives the information released by other nodes and calculates the distance between them. Then node \( i \) selects the four nearest nodes \( N_a, N_b, N_c, N_d \) and stores the information of these nodes.
4. Node \( i \) calculates \( Hash_i = Hash(N_a||N_b||N_c||N_d) \) and send to the four nodes \( \{Hash_i, H(ID_i)\} \).
5. The remaining four nodes receive the message send by node \( i \) and return the same message after authentication.

Group is formed and nodes within the group trust each other after the process above. When a node in this group dies, the remaining four nodes in the group verify the hash value with each other, and choose a new routing path according to the distance between the nodes. A new intra-group chain is built and continue the data transmission. Compared with the traditional re-chain of the entire network, PEGASIS-GAMT reflects a better energy consumption reduction and a lower complexity as an extended protocol.

4. Simulation Results

The parameter used in this experiment and the values of the parameters is given in table 1. In order to demonstrate the fairness of the simulation experiment, the comparison experiments in this paper are all based on the same node distribution and parameter setting. Before the death of the first node, the
energy consumption generated by PEGASIS and PEGASIS-GAMT is exactly the same. Only when the first dead node appears, PEGASIS-GAMT will show its additional advantages.

Table 1. Parameter description and values

| Parameter  | Type                                      | Value                  |
|------------|-------------------------------------------|------------------------|
| $E_{\text{init}}$ | The initial energy of sensor node          | 0.25 J                 |
| $E_{\text{elec}}$ | The electronics energy                    | 50 nJ/bit              |
| $E_{d}$     | The energy for data aggregation           | 5 nJ/bit/signal        |
| $\varepsilon_{\text{amp}}$ | The communication energy parameter     | 100 pJ/bit/m$^2$       |
| $l$         | Packet length                             | 2000 bit               |
| $l_a$       | Authentication packet length              | 50 bit                 |
| $H$         | The height of simulation area             | 50 m                   |
| $L$         | The length of simulation area             | 10 m                   |
| $N$         | The number of wireless sensor nodes       | 30                     |
| $r$         | Number of rounds                          | 1500                   |

Figure 5 shows the residual energy of all nodes in per round of the network. Nodes begin to die at the 970th round, then the traditional PEGASIS begin to re-chain after the node died, which increased the energy consumption of the whole network. The red line representing the residual energy of PEGASIS shows an accelerating downward trend. When the node dies, PEGASIS-GAMT replaces the communication cost of the whole network with a small amount of communication cost in the group by means of re-chain within the group, which can undoubtedly save energy consumption for the network. As shown in the blue line representing PEGASIS-GAMT in the figure, the remaining energy decreases slowly.

![Figure 5. Total residual energy per round](image1)

![Figure 6. Number of nodes alive per round](image2)

As shown in figure 6, the number of nodes alive per round corresponds to the content shown in figure 5. The nodes of PEGASIS are all dead at the 1069th round. The blue line representing PEGASIS-GAMT, which is obviously later then the red line, the nodes of it are all dead at the 1122th round, which is 53 rounds later than PEGASIS protocol. Considering when the first node dies, all the nodes in the network have little residual energy at the time, the 53 rounds longer network life is uneasiness.

5. Conclusions

This paper has proposed an extended PEGASIS protocol based group authentication with Merkle Tree. The extension lies in the fact that there is no change to the existing algorithm, but the group scheme and intra-group re-chain are added. In a newly deployed network, the proposed scheme only groups the nodes according to the optimal value and form intra-group mutual trust. The steady phase from existing PEGASIS are exactly the same. The extension reflects the difference through intra-group re-chain instead of re-chain the entire network, only when the node starts to die. The simulation results show that PEGASIS-GAMT not only prolongs the network life dramatically, but also avoids the attacker malicious deception by false information. In the future, this extended scheme has strong universality and can be ported to other existing CWSN protocols.
Acknowledgments
The work presented in this paper was supported in part by the NSF of China with Grant 61562038.

References
[1] Zhou, G., Zhu, Z., Chen, G., & Hu, N. (2009, May). Energy-efficient chain-type wireless sensor network for gas monitoring. In 2009 Second International Conference on Information and Computing Science (Vol. 2, pp. 125-128). IEEE.
[2] Ok, C. S., Lee, S., Mitra, P., & Kumara, S. (2009). Distributed energy balanced routing for wireless sensor networks. Computers & Industrial Engineering, 57(1), 125-135.
[3] Jung, S. M., Han, Y. J., & Chung, T. M. (2007, February). The concentric clustering scheme for efficient energy consumption in the PEGASIS. In The 9th international conference on advanced communication technology (Vol. 1, pp. 260-265). IEEE.
[4] Mishra, A. K., Rahman, R. U., Bharadwaj, R., & Sharma, R. (2015, October). An enhancement of PEGASIS protocol with improved network lifetime for Wireless Sensor Networks. In 2015 IEEE Power, Communication and Information Technology Conference (PCITC) (pp. 142-147). IEEE.
[5] Ramluckun, N., & Bassoo, V. (2018). Energy-efficient chain-cluster based intelligent routing technique for Wireless Sensor Networks. Applied Computing and Informatics.
[6] Danzi, P., Kalor, A. E., Stefanović, Č., & Popovski, P. (2019). Delay and Communication Tradeoffs for Blockchain Systems with Lightweight IoT Clients. IEEE Internet of Things Journal, 6(2), 2354-2365.
[7] Xu, L., Chen, L., Gao, Z., Xu, S., & Shi, W. (2018). Efficient Public Blockchain Client for Lightweight Users. arXiv preprint arXiv:1811.04900.
[8] Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000, January). Energy-efficient communication protocol for wireless microsensor networks. In Proceedings of the 33rd annual Hawaii international conference on system sciences (pp. 10-pp). IEEE.
[9] Alrawais, A., Alhothaily, A., Mei, B., Song, T., & Cheng, X. (2018). An efficient revocation scheme for vehicular ad-hoc networks. Procedia Computer Science, 129, 312-318.