Reversible Authentication of Wireless Sensor Network Based on Prediction-Error Histogram and CRC

Guangyong Gao(✉), Caixue Zhou, Zongmin Cui, Shimao Yao, and Zhijun Chong

School of Information Science & Technology,
Jiujiang University, Jiujiang 332005, China
gaoguangyong@163.com

Abstract. In this paper, a reversible authentication scheme for wireless sensor network (WSN) is proposed. Firstly, the WSN data stream is divided into some authentication groups, and each authentication group is composed of a generator group and a carrier group. Then the cyclical redundancy check (CRC) code of generator group is produced as the authentication information. In the carrier group, using the prediction-error-based histogram shifting algorithm, the authentication information is reversibly embedded into the fluctuation region of prediction-error histogram (PEH), not the smooth region. Experimental results and analysis demonstrate that compared with previous schemes, the proposed scheme achieves better performances on computation complex, false tampering alarm and attraction to attackers.

Keywords: Wireless sensor network (WSN) · Reversible authentication · Prediction-error histogram (PEH)

1 Introduction

The data integrity authentication is a core issue of wireless sensor network (WSN). Traditionally, encryption technology is applied to maintaining the security of WSN [1, 2]. But due to the limitation of sensor node resource, it is not fit to apply encryption technology to WSN.

Later, the information hiding technology applied to image security is used for the data authentication of WSN. Compared to encryption, the information hiding is more lightweight on resource consumption. In [3], an information hiding scheme is firstly proposed. The scheme utilize the property of micro-errors of sensor data to embed information, and so long as the introduced error is within a limited scope, the WSN data can be used normally. In [4], a fragile chain-watermarking scheme is developed. This scheme regarded the WSN data as a stream chained by many groups, and generate watermarking by the repeated use of hash function for different groups. The experimental results indicated the fragile chain-watermarking scheme can verify the integrity of WSN data. It is noted that above information hiding schemes is not reversible.

F. Chen and Y. Luo (Eds.): Industrial IoT 2017, LNICST 202, pp. 130–139, 2017.
DOI: 10.1007/978-3-319-60753-5_14

In some special applications such as medical and military fields, the authenticated data is not allowed to be modified. Therefore, for these applications, only the reversible information hiding technology [5, 6] is adopted. In [7], Shi et al. proposed a reversible authentication scheme based on group strategy. Each authentication group is determined dynamically by synchronization points, and the data stream can be totally recovered after the authentication information composed of hash value of authentication group is extracted. In [8], a reversible authentication scheme is achieved based on the cyclical redundancy check (CRC) and odd-even invariability, but this scheme has a high false tampering-alarm rate.

In this paper, a novel reversible authentication scheme for WSN is proposed. The authentication information is generated by calculating the CRC code of generator group. In the fluctuation region of prediction-error histogram (PEH) of carrier group, the authentication information is reversibly embedded by applying the PEH shifting algorithm. Experimental results and analysis demonstrate the proposed scheme achieves better performances than previous reversible authentication schemes.

2 Proposed Scheme

2.1 System Model

A simple WSN system model is provided in Fig. 1, where there are three types of nodes such as sensor node, transmission node and convergence node. The sensor node is responsible to collect periodically the environmental data and embed the authentication information. The embedded WSN data is then sent to convergence node by transmission node. The convergence node, as the central node of WSN, has rich computation resources and strong energy, where the data group is authenticated with extracted embedded-information.

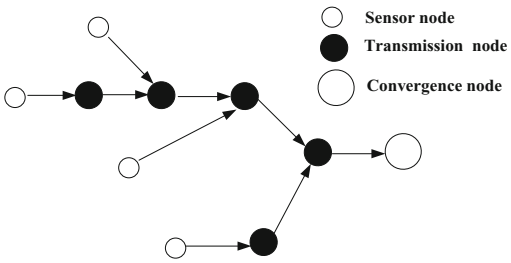

Fig. 1. Diagram of WSN data transmission

In proposed scheme, the WSN data stream collected by sensor node is denoted by V, which is divided into a series of authentication groups. An authentication group consists of a generator group and a carrier group that is shown in Fig. 2. Each element in an authentication group is presented by v_i. The CRC code of each generator group, as authentication information, is embedded reversibly into the corresponding carrier group

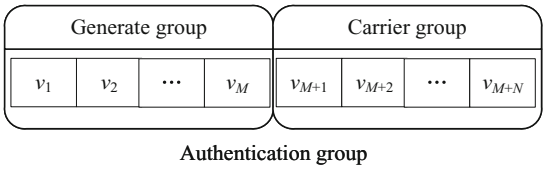

Fig. 2. An example of authentication group

using the prediction error-based histogram shifting algorithm [9]. In the decoding end, the integrity of each authentication group is judged in terms of the comparison between the extracted hidden-information from each carrier group and the CRC code of the corresponding generator group.

2.2 Prediction Error-Based Histogram Shifting Algorithm

2.2.1 Encoding Phase

Let v_i be an element of a carrier group, v_{i+1} *and* v_{i+2} are two neighbor elements of v_i. The prediction value of v_i is denoted by $\hat{v}_i$, which is calculated using Eq. (1).

$$\hat{v}_i = \left\lfloor \frac{v_{i+1} + v_{i+2}}{2} \right\rfloor \tag{1}$$

Then the prediction error between v_i and $\hat{v}_i$, denoted by e_i, is counted using Eq. (2).

$$e_i = v_i - \hat{v}_i \tag{2}$$

An example of PEH is shown in Fig. 3, where the middle part of PEH is smooth region, the two sides of PEH is called as fluctuation region. If the authentication information is embedded into the smooth region, then the smooth property of this region will be destroyed, which may attract the attention of attackers. In contrast, embedding the authentication information into the fluctuation region can not change its fluctuation property. Therefore, the fluctuation region is a better choice to embed the authentication information than the smooth region.

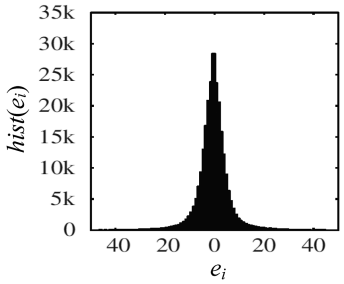

Fig. 3. An example of the prediction-error histogram

Since the prediction errors e_is at two sides of PEH are different, an initial parameter T_m with lower absolute value is selected by Eq. (3).

$$T_m = \min(|\min(e_i)|, \max(e_i)) \tag{3}$$

Then the end parameter T_p is selected in terms of Eq. (4), which should satisfy the capacity demand of embedded authentication information.

$$\begin{cases} \text{maximize } T_p \in (0, 1, 2, \ldots, T_m) \\ \textit{subject to } (\sum\limits_{E=-T_m}^{-T_p} \textit{hist}(E) + \sum\limits_{E=T_p}^{T_m} \textit{hist}(E)) > \textit{capacity} \end{cases} \tag{4}$$

In Eq. (4), *hist*(E) denotes the pixel number with the prediction error of E, and *capacity* means the bit number of the embedded authentication information.

A bit of authentication information, denoted by b, is embedded into the fluctuation region of PEH by Eq. (5).

$$D_i = \begin{cases} e_i + 1, & e_i > E \\ e_i - 1, & e_i < -E \\ e_i + b, & e_i = E \\ e_i - b, & e_i = -E \\ e_i, & else \end{cases} \tag{5}$$

where D_i is the modified prediction error. It is observed from Eq. (5) that the prediction errors larger than E or less than $-E$ are shifted by increasing 1 or decreasing 1, and the prediction errors in the smooth region remain unchanged.

Finally, the marked element value V_i is calculated using Eq. (6).

$$V_i = D_i + \hat{v}_i \tag{6}$$

It is noted that the prediction errors of the neighbor elements v_{i+1}, v_{i+2} *and* v_{i+3} of v_i are not counted and not used to embed information. In addition, in order to recover the original authentication group in the receiving end, the parameters T_m and T_p need to be embedded into the first thirty LSBs of carrier group. Together with the authentication information, the first thirty LSBs of carrier group are embedded the fluctuation region of PEH by Eq. (5).

2.2.2 Decoding Phase

In the receiving end, firstly, the first thirty LSBs of carrier group are extracted to obtain the parameters T_m and T_p. Then the modified prediction error D_i is calculated using Eq. (7).

$$D_i = V_i - \hat{v}_i \tag{7}$$

The hidden information is extracted using Eq. (8), where f_i changes from 0 to $f_{\max} - 1$, H_i changes from T_p to T_m, and $f_{\max} = T_m - T_p + 1$.

$$b = \begin{cases} 0, & D_i = H_i + f_i \quad or\ D_i = -H_i - f_i \\ 1, & D_i = H_i + f_i + 1 \ \ or\ D_i = -H_i - f_i - 1 \end{cases} \tag{8}$$

The original prediction error e_i' is recovered using Eq. (9)

$$e_i' = \begin{cases} D_i - f_i, & D_i = H_i + f_i \\ D_i + f_i, & D_i = -H_i - f_i \\ D_i - f_i - 1, & D_i = H_i + f_i + 1 \\ D_i + f_i + 1, & D_i = -H_i - f_i - 1 \\ D_i, & D_i > -H_i - f_i \ and\ D_i < H_i + f_i \end{cases} \tag{9}$$

The original prediction error larger than T_m is recovered using Eq. (10).

$$e_i' = \begin{cases} D_i - f_{\max}, & D_i > T_m + f_{\max} \\ -D_i + f_{\max}, & D_i < -T_m - f_{\max i} \end{cases} \tag{10}$$

The original WSN data element is restored by Eq. (11).

$$v_i = \hat{v}_i + e_i' \tag{11}$$

Finally, the first thirty LSBs of carrier group extracted by Eq. (8) is written back, so that all elements of a carrier group are completely restored.

2.3 Generation of Authentication Information

The CRC code of each generator group is calculated as authentication information using Algorithm 1.

Algorithm 1: Generation of authentication information

Input: Generator group element v_i
Output: CRC code
Step1. The generator group element v_i is translated to binary expression, which is denoted by bv_i, then bv_i is represented as a polynomial. For example, '1011001' is expressed as $x^6 + x^4 + x^3 + 1$.
Step 2. bv_i is shifted left l bits and regarded as dividend, where l is a parameter controlling the length of CRC code. The polynomial $x^l + x^{l-2} + x^{l-3} + 1$ is taken as divisor.
Step 3. The module-2 division between the dividend and divisor polynomials is used to obtain l-bits remainder cv_i, which is CRC code of v_i.
Step 4. XOR operation among the CRC codes of all elements of current generator group is performed to generate the authentication information W, as shown in Eq. (12) where $\oplus$ indicates XOR operation.

$$W = cv_1 \oplus cv_2 \oplus \cdots \oplus cv_{M\text{-}1} \oplus cv_M \tag{12}$$

2.4 Embedding of Authentication Information

From Sects. 2.2 and 2.3, it is known that the information to be embedded is composed of the authentication information W and the first thirty LSBs of carrier group with a length of $(l + 30)$ bits. In order to keep the description compact, we still adopt W to represent all information to be embedded, and call W as authentication information. To assure the total embedding of W, the length of carrier group is set as $(l + 30) \times 9 + 30$. Moreover, for the convenience to detect carrier group in receiving end, an element as a string "000000" is added as last element of carrier group, and the notation *EOC* is used to denote the element. Therefore, each carrier group includes $(l + 30) \times 9 + 31$ elements. The embedding procedure of authentication information W may refer to Algorithm 2.

Algorithm 2: Embedding

```
while (stream V is not over) do //V is the WSN data stream
    while (stream V is not over, and i ≤ M) do
        Buffer(v_i); //buffer current data element
        calculate cv_i; // get CRC code of v_i according to Algorithm 1
        CRC = CRC ⊕ cv_i; //calculate CRC code of current generator group
    end while
    if (stream V is not over) do
        Buffer(carrier group); // buffer (l + 30)×9 + 30 elements as current carrier group
        W = CRC ⋃ LSBs //W includes the first thirty LSBs of carrier group
        embed W according to the embedding method given in Section 2.2.1
        Insert EOC as last element of carrier group
    end if
end while
```

2.5 Information Extraction and Authentication

In receiving end, the receiver can extract the embedded authentication information, and confirm if a transmitted data group is tampered in terms of the comparison between the generated authentication information by generator group and the extracted authentication information. If a group is intact, then the group will be restored to its original state. The concrete procedure may refer to Algorithm 3.

Algorithm 3: Extraction and authentication

```
while (stream V is not over) do // V is the WSN data stream
  buffer(G); //buffer current authentication group
  while (group G is not over, and i ≤ M) do
    buffer(v_i);
    calculate cv_i; // get CRC code of v_i according to Algorithm 1
    CRC = CRC ⊕ cv_i; //calculate CRC code of current generator group
  end while
  if (group G is not over) do
    buffer(carrier group);
    extract W according to the information extraction method given in Section 2.2.2;
    C R Cis obtained by separating W;
    if (CRC == CRC′) do
      G passed the authentication;
      restore G according to the recovery method given in Section 2.2.2;
    else
      authentication on G failed;
    end if
  end if
end while
```

3 Experimental Results and Analysis

3.1 Experimental Data

The original WSN data in the experiments is from the real WSN deployed in Intel Berkeley Lab [10]. In this network, a series of sensor nodes collect periodically four types of data such as temperature, humidity, light and voltage. The experimental data used in each experiment consists of the information of 10000 times collection by a node, i.e., each experimental data stream includes 10000 elements. As shown in Fig. 4,

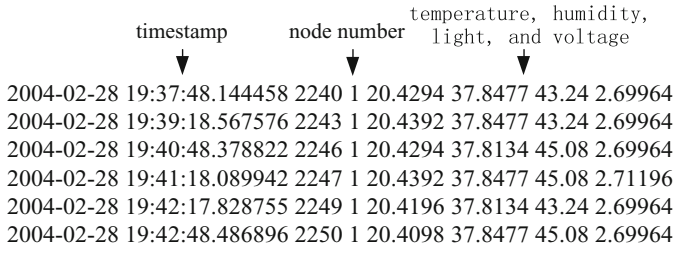

Fig. 4. Original WSN data from Intel Berkeley Lab

every element is composed of eight columns data, involving time stamp, node number and four types of collected data. For simplicity, we only consider to authenticate the integrity of four types of collected data. The verification experiments for the performance of proposed scheme are performed by MATLAB simulation.

3.2 Integrity Authentication Test

The integrity authentication tests for six common tampering attacks are performed and the test results are listed in Table 1. The six attacks involve randomly inserting, deleting and modifying an element or several elements in generator group and carrier group. If the tampering attack is not detected, then the integrity authentication is failing. Each kind of attack adopts 100 random samples. It is indicated from Table 1 that the unsuccessful detection number is zero for all kinds of attacks, and the successful authentication rate is 100%, which demonstrate the proposed authentication scheme is reliable.

Table 1. Integrity authentication test

Tampering type	Number of detection failure	Authentication succeed rate
Insert an element or several elements in generator group	0	100%
Insert an element or several elements in carrier group	0	100%
Delete an element or several elements in generator group	0	100%
Delete an element or several elements in carrier group	0	100%
Modify an element or several elements in generator group	0	100%
Modify an element or several elements in carrier group	0	100%
Average	0	100%

Furthermore, for the authentication group not being tampered, the comparison test between the recovered group after extracting the embedded information and the corresponding original group shows they are consistent, namely, the authentication group can be completely restored.

3.3 Complexity Analysis

Firstly, the space complexity is analyzed. In the sensor node, the buffer size containing carrier group at the most is needed to finish the embedding of authentication information. Comparison with the buffer size containing total authentication group

demanded by the traditional WSN authentication methods, the proposed method has some advantages. In the convergence node, the buffer size containing total current authentication group is needed to achieve information extraction and authentication. Since the convergence node has rich resource, the demand for buffer size is easy to satisfy. For the side information, only two additional parameters, T_m and T_p, are needed to transmit.

Next, the analysis for time complexity is given. In the proposed scheme, the time consumption mainly involves the group division of data stream, the calculation of CRC code, and the embedding and extraction of authentication information. Assume the length of data stream is L, then the time complexity for the group division of data stream is $O(L)$. The time consumption for the calculation of CRC code is closely related with the number of generator group elements M and the controlling parameter of CRC code length l, so the time complexity is $O(M \times l)$. Moreover, in terms of the Algorithms 2 and 3, it is known that the time complexities for embedding and extraction are $O(l \times L)$ and $O(M \times l) + O(L)$, respectively. Therefore, the total time complexity is $O(l \times L)$.

3.4 Detection Efficiency

The false negative rate is adopted to measure the detection efficiency of the proposed scheme, and is analyzed under three kinds of attacks such as inserting, deleting and modifying an element. Assuming inserting an element in generator group, the position probability of this inserted element is $1/M$. Meanwhile, the probability to let the CRC code after inserting be consistent with that before inserting is $1/2^l$. Therefore, the false negative rate is $1/(M \times 2^l)$. If an element is inserted in carrier group, then the analysis for the false negative rate is more complex. But surely the false negative rate in this case is less than $1/(M \times 2^l)$. For other two attacks such as deleting and modifying an element, the false negative rates are similar to that of inserting attack. It is concluded from the above analysis that the probability to make a false judgement for tampering attack will be very low if the values of M and l are big enough.

3.5 Comparisons Among Several Schemes

The comparisons among Shi's scheme [7], Wu's scheme [8] and the proposed scheme are conducted. In Shi's scheme, the hash values of generator group are used as authentication information. In comparison to CRC code, the computation complex of hash function is higher. In Wu's scheme, an authentication group insists of current data group and former data group, and the sizes of the current data group and former data group are fixed, i.e., the division of authentication group is based on fixed length. So when an element is inserted, all authentication groups will be changed so that the false tampering alarm may be induced. In the proposed scheme, the authentication group is divided according to the end notation, hence the situation similar to Wu's scheme does not occur. In addition, Shi's scheme and Wu's scheme do not distinguish the smooth and fluctuation regions when the authentication information is embedded. However, the

proposed scheme only embeds the authentication information into the fluctuation region, which may attract less the attention of attackers.

4 Conclusion

Based on CRC code and prediction-error histogram, this paper proposes a reversible authentication scheme for WSN. For decreasing the computation complexity, the CRC code replacing the hash function is adopted to produce the authentication information. To increase the imperceptibility of embedding information into the carrier group, the fluctuation region of prediction-error histogram is used to hide the authentication information. In the future work, degrading further the time complexity will be considered.

Acknowledgments. This work is supported in part by the National Natural Science Foundation of China (Grant No. 61662039, 61362032, 61462048), the Natural Science Foundation of Jiangxi Province, China (Grant No. 20171BAB202004, 20151BAB207003, 20161BAB202036), and the State Scholarship for Overseas Studies (Grant No. 201408360019).

References

1. Seo, S.H., Won, J., Sultana, S., Bertino, E.: Effective key management in dynamic wireless sensor networks. IEEE Trans. Inform. Forensics Secur. **10**(2), 371–383 (2015)
2. Marzi H.: A security model for wireless sensor networks. In: IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, pp. 64–69. IEEE Press (2014)
3. Wong J.L, Feng J., Kirovski D.: Security in sensor networks: watermarking techniques. In: Wireless Sensor Networks, pp. 305–323 (2004)
4. Guo, H., Li, Y., Jajodia, S.: Chaining watermarks for detecting malicious modifications to streaming data. Inform. Sci. **177**(1), 281–298 (2007)
5. Thodi D.M., Rodriquez J.J.: Prediction-error-based reversible watermarking, In: IEEE International Conference on Image Processing, pp. 1549–1552. IEEE Press (2004)
6. Li, X., Zhang, W., Gui, X.: Efficient reversible data hiding based on multiple histograms modification. IEEE Trans. Inform. Forensics Secur. **10**(9), 2016–2027 (2015)
7. Shi, X., Xiao, D.: A reversible watermarking authentication scheme for wireless sensor networks. Inform. Sci. **240**(10), 173–183 (2013)
8. Wu, H., Chen, Y., Ji, Z.: Wireless sensor networks authentication algorithm based on CRC and reversible digital watermarking. Comput. Appl. Softw. **33**(6), 294–298 (2016)
9. Yang, Y., Zhang, W., Hou, D., Wang, H.: Research and prospect of reversible data hiding method with contrast enhancement. Chin. J. Net. Inform. Secur. **2**(4), 12–19 (2016)
10. http://db.lcs.mit.edu/labdata/labdata.html