International Journal of Grid Distribution Computing
Vol.8, No.2 (2015), pp.143-154
http://dx.doi.org/10.14257/ijgdc.2015.8.2.14

A Reversible Watermarking Scheme Based on Difference Expansion for Wireless Sensor Networks

Qun Ding[1,2] , Baowei Wang[1,2] , Xingming Sun[1,2], Jinwei Wang[1,2] and Jian Shen[1,2]

[1]*School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China*
[2]*Jiangsu Engineering Center of Network Monitoring, Nanjing 210044, China*
dingqun0926@163.com, wbw.first@163.com, sunnudt@163.com, wjwei_2004@163.com, s_shenjian@126.com

Abstract

Wireless sensor networks are self-organized and data-centric. Sensors are typically operated in uncontrolled and possibly hostile environment. Watermark technology has been used for integrity and copyright protection. Some specific application scenarios, such as military, telemedicine and satellite remote sensing, are strict with data integrity and accuracy. They do not allow any change and must be non-destructive data recovery after the requirements of certification. However, previous watermarking-based approaches always make irreversible modifications in the watermarked data. In this paper, we propose a novel reversible watermarking authentication scheme RDE for WSNs, which is based on the difference expansion of a generalized integer transform. Source sensors use a one-way hash function to generate watermark information depending on the adjacent data, and then embed it into these data. After receiving data, the manage node restores the original data and verifies the reliability. The real experiments also demonstrate that our algorithm can losslessly authenticate sensor data with low energy overhead.

Keywords: *Wireless sensor networks, reversible watermarking, data authentication, difference expansion*

1. Introduction

Wireless sensor network (WSN) is a self-organized and data-centric multi-hop network composed of base stations and a large number of multifunctional sensor nodes. Small-sized sensors are scattered in unattended areas, which have limited computing capabilities, communicate wirelessly, and operate on batteries with limited power. Data from the sensor field is collected and sent to a base station, which usually has unlimited resource compared to sensor nodes. WSN is typically used in military, environmental monitoring, smart home, traffic management and other fields. Data integrity is a core requirement for secure sensor data in WSN. But due to the limited computational capacity, storage space and energy, traditional solutions for integrity based on cryptography are often too expensive for sensor nodes. So the digital watermarking technology is introduced. Watermarking has been popularly used in security scheme in WSNs [1-6], mainly based on changing the least significant bit or embedding some blank space, which increases the length of the data packet or destructs the accuracy of the original data to some extent. Such changes often do not affect the visual perception and information expression of images or any other multimedia data with high redundancy. However, it is unacceptable in some critical applications such as military applications or medical care, which require the absolutely accurate original data.

In this paper, we proposed a novel reversible authentication scheme based on lossless

ISSN: 2005-4262 IJGDC

fragile watermarking algorithm for WSNs. The proposed RDE scheme can verify the sensory data through the embedded watermark bits, and restore the original data completely. It meets the demand of the original data in some specific applications in WSNs with a small amount of additional overhead. What's more, unlike previous algorithms realized in simulation system, our scheme is testified in a real wireless sensor network environment. The rest of the paper is organized as follows. Section 2 presents an overview of related work. Section 3 provides the detailed description of our proposed method. The experimental results and the performance evaluation are discussed in Section 4. Finally, Section 5 concludes this paper.

2. Related Work

In [1], J. Feng, *et al*., first proposed the method of applying watermarking into WSN, which was used for copyright protection of the sensor data. The key idea is to impose additional constraints during the data collection or sensor data processing. The method was based on trilateration localization, allowing the distance between expectations and measured value to have a certain error. The algorithm can apply to the real-time requirements in WSN and not affect the normal use of perception data as long as the error introduced by the embedded hidden information is in a limited range. In [2], Guo, *et al*., proposed a novel fragile watermarking algorithm which verifies the integrity of streaming data at the application layer. The data are divided into groups based on synchronization points, so each group can be synchronized and any modifications made to one group only affect no more than two groups. Watermarks are chained across groups so that no matter how many data are deleted, these deletions can be correctly detected. This scheme can efficiently detect and locate modifications and ensure the completeness of data streams. In [3], Xiangrong Xiao, *et al*., proposed a novel secure transmission strategy based on information hiding (IH). They acquire sensitive data security using the advantage of IH technique without encryption. Sensor nodes collect two groups of data: sensitive data and ordinary data, then forward the ordinary data with sensitive data embedded.

However, most of these watermarking-based approaches always make irreversible modifications in the watermarked data, which affect the accuracy of the sensor data and increase the length of the data packet.

Reversible watermarking is a new research orientation on watermarking certification in the field of multimedia digital image. Reversible watermarking algorithm can not only complete the accurate certification of image, but restore the original image completely after extraction of the embedded data. Tian [7] first proposed a classical reversible watermarking algorithm for gray-scale images. He used a difference expansion transform of a pair of pixels to devise a high-capacity and low-distortion reversible watermark. His algorithm divides the image into pairs of pixels, then he embeds one bit into the difference of the pixels of each pair from those pairs that are not expected to cause an overflow or underflow. In a single pass, Tian's algorithm is capable of embedding as high as 1/2 bits/pixel. Adnan M. Alattar [8] extend Tian's algorithm using difference expansion of vectors, instead of pairs, to improve the hiding ability and the computation efficiency of the algorithm. This approach allows the algorithm to embed several bits in every vector in a single pass through the image data. The embedding capacity of Alattar's algorithm can be raised to N-1/N bits/pixel. The thought of difference expansion algorithm is exploiting the correlation among the pixels. Similarly, the data element sensed by sensor nodes can be considered as one pixel of the image, which motivates us to adapt its reversible idea for data authentication in WSNs.

3. The Proposed Solution RDE

The security model of WSN based on reversible watermarking is illustrated in Figure 1. The wireless sensor networks are composed of the sensor node, sink node, internet or satellite and manage node. The sensor node is the encoder in which the watermark bits are generated and embedded. The intermediate nodes transfer the watermarked information to the sink node through multiple hops routing. The manage node is the decoder in which the watermark bits are extracted and the original data is restored. To ensure the reliability of data transmission between the two sides of the communication and verify the authenticity of the data, we embed watermarking into sensory data stream using difference expansion of a generalized integer transform .The scheme can restore the original data completely, which satisfy the invertible demand of some specific high accuracy application environments.

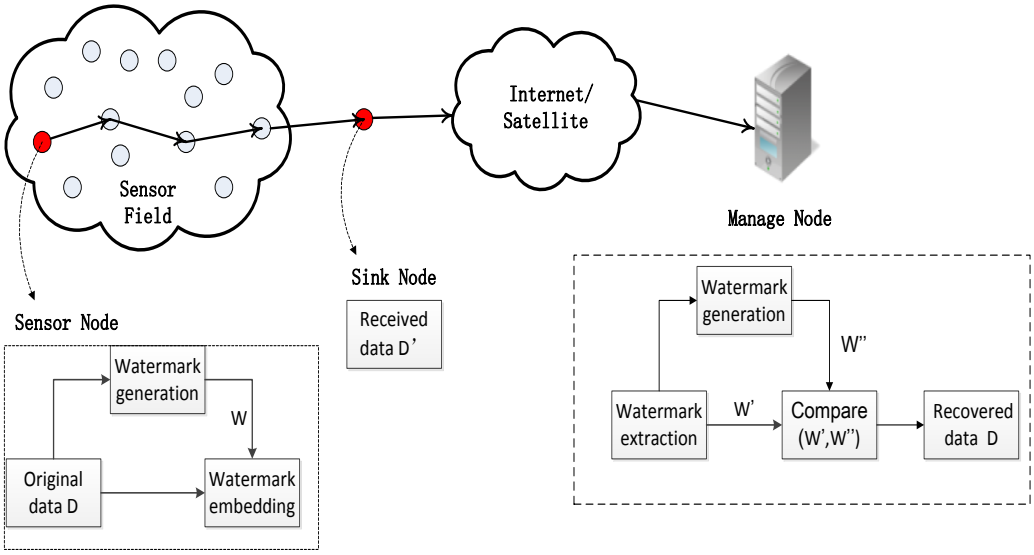

Figure 1. RDE Security Model

3.1. Notations and Parameters

Definition: In RDE, the data packet is represented as packet=(system-data, send-data), where system-data is predefined by the system, including such fixed information as route, packet length, *etc.*,; send-data is the most important part of the packet, collected by sensor nodes. The original send-data D is defined as $D = (S_0, S_1 \cdots S_{n-1})$ where $S_i = (t, d_0, d_1 \cdots d_{m-1})$ denotes the data collected in the i^{th} working period $i = (0,1 \cdots n-1)$; t is the time stamp of the i^{th} collected data; $d_j (j = 0,1 \cdots m-1)$ denotes the collected data item acquired during one working period.

Rule1: watermark generation

The watermark information is generated depended on the original data itself. We denote HASH as a cryptographic hash function such as MD5 or SHA, which takes a message of arbitrary length and secret key K, and produces an output of fixed length. For each item S_i, key K and its time stamp were input into a one-way hash function Hash and $H_i = Hash(K, S_i.t)$. The watermarking information W is a group hash value which is computed as the hash of the concatenation of all individual hash values of data elements in the group and composed of n-1 bits binary data. W is denoted as $W = GroupXOR(H_0 \oplus H_1 \oplus \cdots \oplus H_{n-1})$, in which "$\oplus$" is the XOR operation.

Rule 2: watermark embedding

After generating the watermarking information W, the process of embedding starts. We need to calculate the weighted average of all the sensed data in a packet which is denoted as u and then compute the differences between $S_i.d_j (i = 1, 2 \cdots n-1)(j = 0, 1 \cdots m-1)$ and $S_0.d_j$ respectively. The difference value is denoted as $S_i.h_j$, computed between each period data and the first period data for the same data element. Then each of the difference is required to have a 1-bit left-shifted operation and one bit of W is embedded into the vacant least significant bit (LSB). Since the same type of data between the adjacent periods is extremely approximate, the difference value changes in a small range. So underflow or overflow caused by the 1-bit left-shifted operation is absolutely impossible. The watermarked data D' can be obtained by the inverse difference expansion transform.

Rule 3: watermark extraction and data recovery

When the base station received data D' at the recipient side, the weighted average of all the received data and the differences between $S_i.d_j'$ and $S_0.d_j'$ are calculated respectively the same as the embedding. And then the watermarking information W' is extracted from the differences. The original data D can be regained according to the invertible transform. Obviously, the watermarking information W'' can be recalculated by the recovered data D using the same rule as the embedding. It is proved that the data is not damaged during the transmission only when W' is equal to W''.

3.2. RDE Embedding Algorithm

The watermark embedding algorithm is shown in **Algorithm 1**. The intermediate nodes just need to forward packets. First the watermarking information W is generated according to rule 1. The original data contains n working periods and during each period m data items are collected. The binary representation of W is $W_{n-1} \cdots W_2 W_1$, and the n-1 bits hash values are the candidates of watermark bits. Then we apply the difference expansion of a generalized integer transform technique to water embedding. For every data element in a data packet, 1 bit watermark is embedded into the least significant bit of the difference value between the current data and the first data. Therefore, n data elements should be embedded into n-1 bits watermark.

Algorithm 1. Digital Watermarking Embedding

Input: The original data D, key K

Output: The watermarked data D'

Steps:

1. for(int i=0; i<n; i++)
2. $H_i = Hash(K, S_i.t)$
3. end for
4. $W = GroupXOR(H_0 \oplus H_1 \oplus \cdots \oplus H_{n-1})$
5. for(int j=0; j<m; j++)
6. $u = \left\lfloor \frac{|S_0.d_j + S_1.d_j + \cdots + S_{n-1}.d_j|}{n} \right\rfloor$
7. for(int i=1; i<n; i++)
8. $S_i.h_j = S_i.d_j - S_0$
9. $S_i.h_j' = 2 \times S_i.h_j + W$

10. $S_0.d_j' = u - \left\lfloor \frac{S_1.h_j' + \cdots + S_{n-1}.h_j'}{n} \right\rfloor$
11. $S_i.d_j' = S_0.d_j' + S_i.h_j'$
12. end for
13. end for
14. Send(D′)

The process of generating watermarking information is listed in line 1-4.Then the average value, together with the difference values are calculated in line 5-8. The difference values represented in binary format are shifted left by one bit and then the watermark bit w is appended into the vacant least significant bit (LSB) in line 9. Actually, the one bit left- shift operation of binary number is equal to the decimal digit multiplied by two. One bit watermarking information is embedded into the vacant LSB at once time. Then the new difference value can be obtained as line 10-11. The weighted average u remains unchanged. At last, the watermarked data is gained according to the inverse transform and then sent to the base station in line 14.

3.3. RDE Extraction and Recovery Algorithm

The algorithm for watermark detection and original data recovery is shown in **Algorithm 2**. The resource and energy of the sink is not constrained strictly by comparison to the sensor. When the watermarked data reach the base station, the difference value and the average are calculated firstly according to rule 3 in line 1-4. The sink extract watermarking information W' from the least significant bit (LSB) of the difference value in line 5 and then shift all the difference value right by one bit in line 6. The original data can be regained by the inverse transform in line 7-8.The watermarking information W'' is recalculated according to the original data in line 11-14. Then the computed bits W'' and the extracted watermark bits W' are verified to decide whether the data integrity has been damaged in line 15-18.

Algorithm 2. Digital Watermarking Extraction and Original Recovery

Input: The watermarked data D', key K
Output: The original data D, and the result of data integrity
Steps:
1. for(int j=0; j<m; j++)
2. for(int i=1; i<n; i++)
3. $u = \left\lfloor \frac{S_0.d_j' + S_1.d_j' + \cdots + S_n.d_j'}{n} \right\rfloor$
4. $S_i.h_j' = S_i.d_j' - S_0.d_j'$
5. $W'[i] = LSB(S_i.h_j')$
6. $S_i.h_j = \left\lfloor \frac{S_i.h_j'}{2} \right\rfloor$
7. $S_0.d_j = u - \left\lfloor \frac{S_1.h_j + \cdots + S_{n-1}.h_1}{n} \right\rfloor$
8. $S_i.d_j = S_0.d_j + i$
9. end for

10. end for
11. for(int i=0; i<n; i++)
12. $H_i = Hash(K, S_i.t)$
13. end for
14. $W'' = GroupXOR(H_0 \oplus H_1 \oplus \cdots \oplus H_{n-1})$
15. If (Compare(W′,W″)==Equal)
16. D marked tampered
17. Else
18. D marked authenticated

4. Performance Evaluation

4.1. Experiment Setup

In this section, we have performed real experiments to evaluate the performance of the proposed watermarking scheme. 20 Telosb nodes shown in Figure 2 were deployed in a real wireless sensor network environment, while a sink node is connected to a computer which is used to receive and calculate the data. We use the TinyOS operating system, IEEE 802.15.4 wireless communication stack and the collection tree routing protocol (CTP). The size of the packet payload is 28 bytes for node ID, temperature and humidity. Temperature, humidity data were gathered every one minute. The data packet was transmitted every three minutes including 12 sense data. The nesC language was used to implete the embedding algorithms at the sender sensor; whereas C# was used for watermark extraction and integrity check algorithms at the base station.

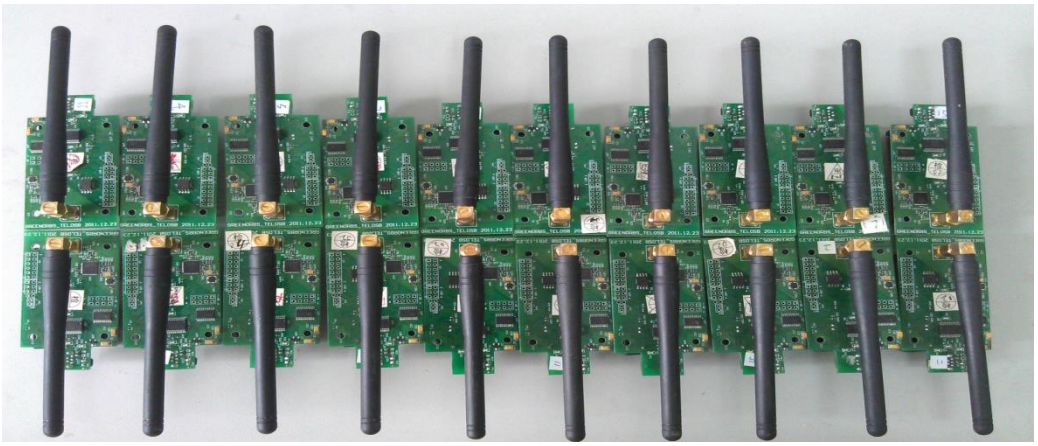

Figure 2. Sensors used in Our Experiments

4.2. Experiment Results

4.2.1. Invisibility and Accuracy:

The numeric comparison among the original data, the embedded data and the restored data at the selected time are shown in Figure 3. Comparing the original data with the embedded data, the watermarking errors are minuscule, since only a least significant bit is changed in the difference value of the adjacent data, which may be ignorable in many applications, because there is a significant similarity among the adjacent data items. Likewise, comparing the original data with the restored data, RDE can regain the original data with scarcely error.

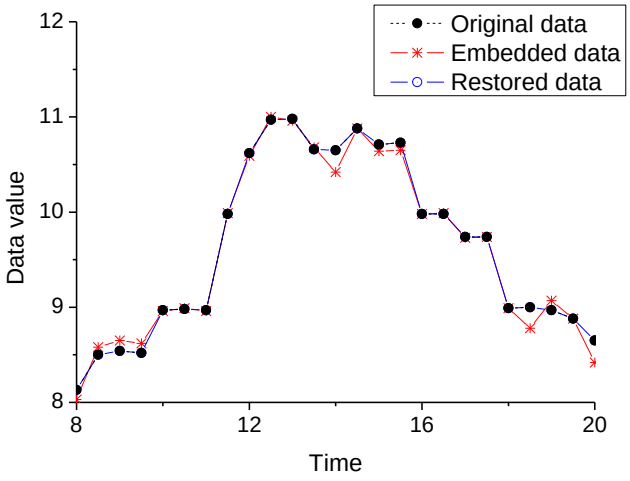

Figure 3. Numerical Comparison Before and After Watermarking Embedding

Table 1 also shows the statistical comparison in terms of the rate of change of the mean and variance of each attribute in the whole stream of tuples. Change of some small percentages in mean and variance indicate that our watermarking scheme have fairly better invisibility and can regain the original data stream successfully.

Table 1. Change in Mean and Variance after Watermark Insertion and Extraction

	Original data		Embedded data				Restored data	
attribute	mean	variance	mean	variance	change in mean	change in variance	mean	variance
1	9.63	0.8747	9.60	0.8917	0.3115	1.9435	9.63	0.8747
2	24.46	0.0004	24.42	0.00025	0.0016	3	24.46	0.0004
3	49.20	0.15	49.13	0.35	0.0014	1.333	49.20	0.15
4	58.32	0.82	58.48	0.80	0.0027	0.024	58.32	0.82
5	23.66	0.0003	23.65	0.00015	0.0004	6.5	23.66	0.0003

4.2.2. Transmission Overhead:

Figure 4 shows that the comparison of the amount of data acquired among the Message Authentication Code (MAC) [9-11] method, add blank character [5], and RDE. The x-axis denotes the packet payload and the y-axis denotes the corresponding amount of data during the transmission.

The MAC method adds the message authentication code directly into the data for transmission, so the amount of data transmission is significantly large. The blank character method embedded spaces selectively according to the corresponding data bit of watermark information is 1 or 0, which increases the length of the data packet and the data cannot be accurately regained. Although the way of the former modifying the least significant bit [12-16] has no effect on the amount of data transmitted, the accuracy of the data has been damaged. RDE scheme not only has no additional transmission overhead but also regains the original data reversibly.

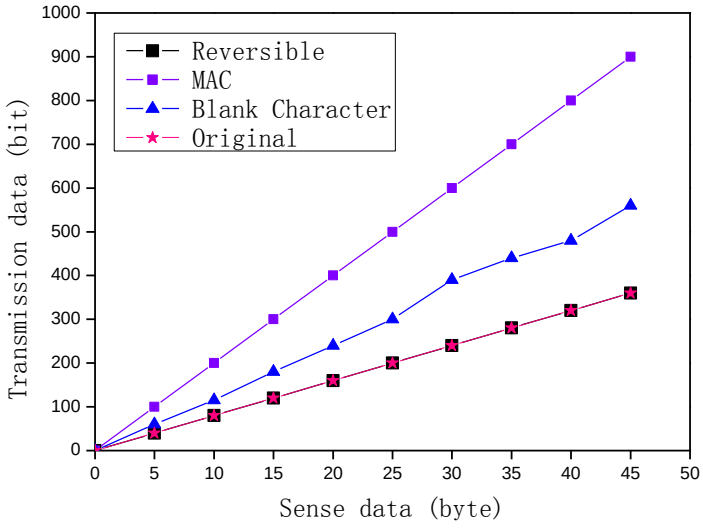

Figure 4. The Amount of Transmission Data

4.2.3. Security:

Attacks are divided into five types: packet forgery attack, selective forwarding, packet replay, packet transfer delay and packet tampering. Select five nodes as the captured nodes and each attack needs to be tested 50 times. The experimental results are presented in Table 2.

Table 2. The Experimental Results of Anti-attack Capability

Attacks	No. of experiments	Success rate (%)
Transfer Delay (2 groups)	50	94
Selective Forwarding	50	100
Packet Forgery	50	100
Packet Tampering	50	100
Packet Forgery	50	100

In our proposed scheme RDE, the watermark information is calculated by the key K and its time stamp. The eavesdropper can't obtain the data information by monitoring communication signal without key. According to the experimental results our method achieved 100% detection on packet forgery attack, selective forwarding, packet replay and packet tampering. It failed three times resisting the packet transfer delay attacks due to the short interval set in the source sensor nodes. Though the transmission delay will cause the delay of watermark extraction, the extracted results are not affected. After the watermarking embedding, the data will be immediately sent to the sink. So the data transmission delay has no influence on the algorithm. Therefore, the experimental results show that our proposed watermarking scheme RDE can effectively detect the integrity of the received information and ensure the authenticity of the data.

4.2.4. Energy Consumption:

When evaluating the energy consumption, we suppose that data storage and forwarding, broadcasting and routing are constant, which is equal to the energy consumption without encryption. So we only need to take into account the extra energy expenditure in watermarking embedding and data transmission in our security scheme proposed in WSN. The calculative overhead increases in terms of watermarking

generation, embedding, extraction and data recovery. We consider arithmetic operations, logical operations, and bitwise operations as lightweight operations. And the notable additional computational overhead is the hash functions whose energy consumption is much less than symmetric encryption and asymmetric encryption [17].

Under normal circumstances, on the order of 3000 instructions can be executed for the energy cost required to transmit one bit over a distance of 100m by radio [18]. Let le denote the 1 instruction energy consumed and the transmission energy consumed will be 3000e. Thus it can be seen that the energy consumption for data transmission is far greater than processing sensor data. In RDE, watermarking information is embedded into the vacant least significant bit of the difference value, which does not occupy additional transmission overhead. Therefore, the energy consumption of the whole wireless network is reduced and the network life cycle is extended.

5. Conclusion

In this paper, a novel reversible watermarking authentication scheme RDE for WSNs has been proposed. Unlike other lossless schemes, our algorithm introduces a new thought that it can restore the original data completely by difference expansion of a generalized integer transform and check the data integrity simultaneously, which is valuable for a wide range of application required high data accuracy. Meanwhile our method does not increase the amount of data transmission. Actual experiments have been conducted in a real deployed wireless sensor network environment. The results have shown not only in terms of data accuracy but also energy consumption and data security the RDE performs manifestly better than other algorithms.

Acknowledgements

This work is supported by the NSFC (61232016, 61173141, 61173142, 61173136, 61103215, 61373132, 61373133), GYHY201206033, 201301030, 2013DFG12860, BC2013012 and PAPD fund.

References

[1] J. Feng and M. Potkonjak, “Real-time watermarking techniques for sensor networks”, SPIE Security and Watermarking of Multimedia Contents, Santa Clara, CA, USA: SPIE Press, **(2003)**, pp. 391-402.
[2] H. Guo, Y. Li and S. Jajodia, “Chaining watermarks for detecting malicious modifications to streaming data”, Information Sciences, no. 177, **(2007)**, pp. 281-298.
[3] X. Xiao and X. Sun, “Secure data transmission of wireless sensor network based on information hiding”, Mobile and Ubiquitous Systems: Networking & Services, **(2007)**, pp. 1-6.
[4] W. Zhang and S. K. Das, “Secure data aggregation in wireless sensor networks: A watermark based authentication supportive approach”, Pervasive and Mobile Computing, **(2008)**, vol. 5, pp. 658-680.
[5] B. Wang, X. Sun and H. Ren, “Multi-mark: Multiple watermarking method for privacy data protection in wireless sensor networks”, Information Technology Journal, **(2011)** October, pp. 833-840.
[6] X., D. X. and Li, “An authentication method for self-nodes based on watermarking in wireless sensor networks”, Wireless Communications, Networking and Mobile Computing, WiCom '09. 5th International Conference, **(2009)**, pp. 1-4.
[7] Tian and Jun, “Reversible data embedding using a difference expansion”, IEEE Transactions on Circuits and System for Video Technology, vol. 13, no. 8, **(2003)**, pp. 890-896.
[8] Alattar and M. Adnan, “Reversible watermark using the difference expansion of a generalized integer transform”, IEEE Transactions on Image Processing, vol. 13, no. 8, **(2004)**, pp. 1147-1156.
[9] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen and D. E. Culler, “SPINS: Security protocols for sensor networks”, Wireless networks, vol. 8, no. 5, **(2002)**, pp. 521-534.
[10] H. Yang, F. Ye, Y. Yuan, S. Lu and W. Arbaugh, “Toward resilient security in wireless sensor networks”, Proceedings of the 6th ACM international symposium on Mobile ad hoc networking and computing. ACM, **(2005)**, pp. 34-45.
[11] L. Hu and D. Evans, “Secure aggregation for wireless networks”, In Workshop on Security and Assurance in Ad hoc Networks, **(2003)** January, pp. 384-391.

[12] R. Sion, M. Atallah and S. Prabhakar, "Resilient rights protection for sensor streams", Proceedings of the Thirtieth international conference on Very large data bases-Volume 30. VLDB Endowment, **(2004)**, pp. 732-743.
[13] W. Zhang and S. K. Das, "Secure data aggregation in wireless sensor networks: A watermark based authentication supportive approach", Pervasive and Mobile Computing, vol. 5, **(2008)**, pp. 658-680.
[14] Juma, H. K., I. Kaya and L, "Watermarking sensor data for protecting the integrity", International Conference on Innovations in Information Technology, **(2008)**, pp. 598-602.
[15] R. Xuejun, "A sensitivity data communication protocol for WSN based on digital watermarking", School of Information and Technology, Northwestern University, Xi'an 710127, China, **(2010)**.
[16] I. Kamel and H. Juma, "A lightweight data integrity scheme for sensor networks", Sensors, **(2011)** November, pp. 4118-4136.
[17] N. R. Potlapally, S. Ravi, A. Raghunathan and N. K. Jha, "Analyzing the energy consumption of security protocols", Proceedings of the 2003 international symposium on Low power electronics and design, ACM, **(2003)**, pp. 30-35.
[18] J. Pottie and W. J. Kaiser, "Wireless integrated network sensors", Communications of the ACM, vol. 43, no. 5, **(2000)**, pp. 51-58.

Authors

Qun Ding, she received her B.S. degree in Software Engineering from Nanjing University of Information Science and Technology, China in 2012. Currently she is studying for her M.S degree in Software Engineering at the same university. Her research interests include wireless networks and network security.

Xingming Sun, he is a professor in the School of Computer and Software, Nanjing University of Information Science and Technology, China from 2011. He received the B.S. degree in Mathematical Science from Hunan Normal University and M.S. degree in Mathematical Science from Dalian University of Technology in 1984 and 1988, respectively. Then, he received the Ph.D. degree in Computer Engineering from Fudan University in 2001. His research interests include information security, network security, cryptography and ubiquitous computing security.

Baowei Wang, he received his B.S. and Ph.D. degrees in Computer Science from Hunan University in 2005 and 2011, respectively. He is currently working as a lecturer in School of Computer and Software, Nanjing University of Information Science and Technology. His research interests include steganography, wireless networks and securing ad hoc networks.

Jinwei Wang, he is an associate professor in the School of Computer and Software, Nanjing University of Information Science and Technology, China from 2010. He received the B.S.degree in Electrical Engineering from Inner Mongol University of Technology in 2000. Then, he received the Ph.D. degree in College of Automation from Nanjing University of Science & Technology in 2007. His research interests include multimedia digital watermark,

multimedia encryption and authentication, cloud data security and multimedia forensics.

Jian Shen, he received the B.E. degree from Nanjing University of Information Science and Technology, Nanjing, China, in 2007 and the M.E. and Ph.D. degrees in Computer Science from Chosun University, Gwangju, Korea, in 2009 and 2012, respectively. Since late 2012, he has been a Professor in the College of Computer and Software at Nanjing University of Information Science & Technology, Nanjing, China. His research interests include computer networking, security systems, mobile computing and networking, ad hoc networks and systems, and ubiquitous sensor networks.