Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.017

IJCSMC, Vol. 7, Issue. 9, September 2018, pg.43 – 48

Encryption Data in Wireless Sensor Network

Solmaz Sharifnia
Department of Mathematics, Shahed University, Tehran, Iran
S_sharifnia@yahoo.com

Abstract— In a paper written by Casteluccia C., Mykletun E. and Tsednik G., an efficient approach is presented for utilizing the aggregation of data in a Wireless Sensor Network and the assuring of end-to-end encryption of data between the leaves and sink. One of the goals of the paper was to minimize the bit transmission between the sensor nodes and therefore to find an efficient encryption algorithm which is simple to implement and in turn would prolong the life of batteries.

Keywords: Encryption data, wireless sensor network, algorithm.

I. Introduction

Wireless Sensor Networks (WSN), by the definition of Holger K. and Willig A. [1], are devices that integrate simple processing power (CPU), memory (storage), and sensing and communication capabilities into a low cost device. WSNs work on the principle of Ad-Hoc networks; each node is a transceiver (it can receive and transmit data.) There are tremendous number of applications of WSN. Some are "Disaster relief applications" [1] where the employed WSN is equipped with thermal sensors measuring the average temperature e.g. in a forest, automatically alarming the fire department if the temperatures get too high. Another field where WSNs could be applied is in military applications, they could measure some important information for the army. In the second case it is obvious that a certain level of security is required; therefore there is a need for encryption algorithms. The need for encryption is always present, especially after the recent incident in which US Army surveillance airplanes, Predator MQ-1's, were not using any encryption algorithms for their surveillance video data [2]. The videos and images of territories under surveillance were monitored by Iraqi militants as well, by just using a shareware windows application, satellite card receiver and a satellite parabola dish [2].
In the following sections first the WSN requirements will be explained, then the encryption using the Boneh-Shaw fingerprinting codes are n-secure codes with ε error and we define new codes that use asymptotically good AG codes concatenated with Boneh-Shaw codes. The error probability of the concatenated construction is $O(1/N) = \exp(-\Omega(\log N))$, with length of order $L = O(c^6 \log c \log N)$, and a decoding algorithm of complexity poly(logN).

II. Wireless Sensor Networks

2.1 WSN Requirements

WSNs have different requirements depending on their application. Usually WSNs are battery powered applications, so one should have in mind that lifetime of a WSN is important and thus the algorithms, which run on them, should be optimized and tested. These algorithms run on tiny Microcontrolling devices (MCU) which are limited in processing and storage space. Since the goal of a WSN is to make a collective conclusion [3], it is important that all sensor nodes work properly and that their lifetime is almost identical for most of them.

2.2 Problems and Issues

A main issue with all wireless devices is their battery power consumption; the more data are being transmitted, the larger the battery consumption is [3]. One way to attack this problem is to reduce the bit transmission. Bit transmission can be reduced by aggregating sensor data [3].
This approach to the problem cannot be applied in all WSNs, but it can be applied where the average, variance, max or min temperature, humidity or some other sensing property is of vital improtance for the WSN.

2.3 Data aggregation as a solution

Aggregating data is a way of compressing the transmitted packet, in a sense that the packet is comprised of only necessary information [3]. Aggregation of data was first introduced in Digital Signal Processing (DSP) applications [4]. There was a need to have an optimal calculation of the average/mean value of all the samples in real time because the DSP Processors did not have enough space in the fast static RAM (SRAM) to store everything. All the sample values of X could not be stored in the SRAM, instead, by summing them (sample values Xn) all up and keeping in mind the number of taken samples (n), it was easy to calculate the average [4].

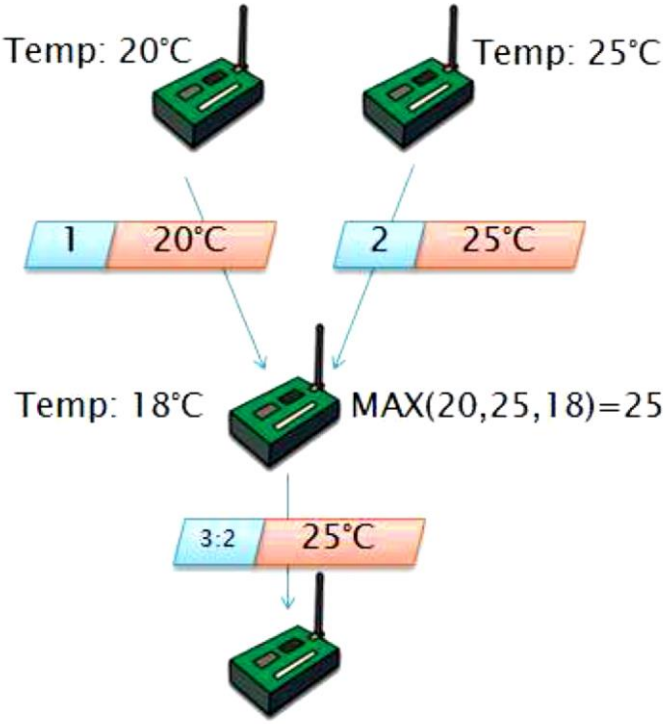

Figure 1: An example of aggregating data. If the node in the middle has access to data, it can choose by itself the maximal temperature value from the three given values and send only the maximal value with the ID of the node whose temperature it is.

Aggregation does not work for every application, i.e. where single reading samples are required e.g. perimeters control [3]. On figure 1, a simple example of data aggregation can be seen (determining the maximal temperature) 1.

III. Encryption of data

The second point in the research paper of Casteluccia C., Mykletun E. and Tsednik G. was encryption of data. Encrypting data is a way of encapsulating the information and protecting it from the outside world, in that sense that nobody should be able to know what information is inside the packet beside the device/person who should receive it. The authors goal was to achieve an end-to-end encryption between the nodes and the sink2.

3.1 End-to-end encryption

In end-to-end encryption no node should be capable of knowing or being able to extract the information from the received packets beside the sink. Using this approach it is possible to guarantee that it will be more difficult for an eavesdropper3 to gain access to the data. Another way of addressing the encryption problem would be to use a global encryption key or only keys between neighboring nodes, but in that case the end-to-end encryption is lost. For the first approach, having one global key, an eavesdropper could gain access to all information by just hacking one node and determining the global key.

3.2 the Boneh-Shaw n-Secure Code

given a subset $X \subseteq \mathrm{IF}_q^n$($\mathrm{IF}_q$is the field of q elements,), we define the *undetectable* positions of X as the components i such that $x_i^r = x_i^s$, $\forall x^r, x^s \in X$, where $x^r = (x_1^n, \dots, x_n^r)$. The undetectable positions form a set denoted by $Z(X)$.

Then we define the *envelope* $\varepsilon(X) \subseteq \mathrm{IF}_q^n$ of X as a set of words than can be derived from X. By the marking assumption, the positions in $Z(X)$ cannot be modified, thus if $y \in \varepsilon(X)$ then $y_i = x_i^r$, $\forall i \in Z(X)$, $\forall x^r \in X$. Here we will consider two envelope definitions, the *narrow-sense* envelope $e(X) = \{ y=(y_1, \dots, y_n) \in (\mathrm{IF}q \cup ?)^n \mid y_i \in \bigcup_{x_i^r} (x_i^r \cup ?)\}$.and the *wide-sense* envelope $E(X) = \{ y = (y_1, \dots, y_n) \in (\mathrm{IF}q \cup ?)^n | y_i = x_i^r, for\ i \in Z(x), x^r \in X\}$.

If $y \in \varepsilon(X)$ then y is a *descendant* of X and any $x \in X$ is a *parent* of y. Note that for the binary case, $q = 2$, these two envelope definitions are equivalent. As we will focus on the binary case, in what follows, we will represent the envelope of X as $\varepsilon(X)$. Moreover, note that since some of the bits in the descendant might be unreadable, then following the convention of Boneh and Shaw in [5], we set these bits to "0" before entering the tracing algorithm.

With the above notation, the fingerprinting problem can be summarized as follows. Let us consider a code C and a c-coalition with fingerprints (codewords) $T = \{t_1, t_2, \dots, t_n\} \subset C$. The coalition creates a new false fingerprint $z \in \varepsilon(T)$ and the distributor D needs to determine which codewords can produce z, that is, D determines a set $G = \{x_1, \dots, x_n\}$ such that $z \in \varepsilon(G)$. If $\emptyset \neq G \subseteq T$ then the code is c-secure.

Boneh and Shaw, in [5] prove that there are no totally c-secure binary codes, that is, any fingerprinting code C, together with an identification algorithm D, it has some error probability, that is, the returned set G can be empty, or some innocent user can be framed, in other words $G \quad T$.

The authors in [5] construct a fingerprinting code n-secure with error probability less than ε. The BS(n, r) code consists of columns of type $c_k = (\overbrace{1 \dots 1}^{k} \overbrace{0 \dots 0}^{l})^T$ for $1 \leq k \leq n-1$, $n = k + l$, where n is the number of users. Moreover each column is repeated r times, generating identical column blocks denoted by C_k^r. If we consider the $n \times r(n-1)$ matrix $C = (C_1^r, \dots, C_{n-1}^r)$, then each row conforms a code word. Then, if each user is unambiguously identified by integer i, where $1 \leq i \leq n$, the scheme assigns codeword (fingerprint/row matrix) i, $1 \leq i \leq n$ to user i. Before embedding the codeword into the digital content a random permutation of the positions is performed.

1.The node images are taken from prof.schindelhauers slides

2- sink is usually the last node in the WSN tree which collects all the data:and in most the cases, it is more powerful than the rest of the node.

3- An eavesdropper is a passive attacker in the middle, who is only listening to the data being transferred between the nodes.

A traitor coalition colludes to create a false fingerprint z, according to the marking assumption. In the identification algorithm, the codewords, of length $r(n-1)$, are divided in $n-1$ blocks, denoted by M_i, $i = 1, \ldots, n-1$, that represent the positions corresponding to the block C_i^r. Taking $w(M_i)$ to be the Hamming weight of block M_i of z, the decoding rules consider user i, $1 < i < n$, one of the traitors if $w(M_i) - w(M_{i-1}) > \lambda_i$, where: $\lambda_i = \sqrt{2(w(M_{i-1}) + w(M_i)) \log(2n/\varepsilon)}$.
Moreover, users 1 and/or n are traitors if $w(M_1) > 0$ and/or $w(M_{n-1}) < r$.
In [5] it is proved that the BS(n, r) code together with this decoding algorithm is n-secure with error probability ε, with a code length of order $O(n^3 \log(n/\varepsilon))$.

IV. Tracing Algorithm

In view of the previous results we can define a tracing algorithm, that given a false fingerprint returns a coalition member with error probability ε. Note that the proposed algorithm only needs to run one time over the bits of the false fingerprint, that is, the complexity time is $O(nc^2 \log(n/\varepsilon))$.
Tracing algorithm:
Input:
– BS(n, r) code with $r > 8\,(c + \sqrt{c-1}\,)^2 log\frac{4n}{\varepsilon}$
– descendant **z** generated by at most c traitors.
Output: List G that contains at least a guilty user with probability 1-ε.

A. ALGORITHM
1. // Initialization:
(a) Set $G = \emptyset$.
(b) Compute $\lambda := \sqrt{2rlog(\frac{4n}{\varepsilon})}$
2. // Identification
(a) **if** $w(M_1) \neq 0$ insert 1 in G.
(b) **if** $w(M_{n-1}) \neq r$ insert n in G.
(c) **for** $i = 2$ to $n-1$
if $w(M_i) - w(M_{i-1}) > 2\lambda$ insert i in G.
3. Output G.

B. Theorem
There exist c-secure fingerprinting codes with N codewords, length $L = O(c^6 \log c \log N)$, and error probability $p_e = O(1/N)$.

Proof. It is well known [6] the existence of families of algebraic-geometric codes (AG), with parameters [n, k, d], over a finite field IF_q, whose parameters asymptotically approach the Tsfasman-Vladut-Zink bound
$\frac{k}{n} \geq 1 - \frac{1}{(\sqrt{q}-1)-\frac{d}{n}}$.
These codes satisfy $n = O(\log N)$, where N is the number of codewords.
Let W be one of the AG codes that approach the Tsfasman-Vladut-Zink bound, with $d > n - n(1-\sigma)/c^2$, where $0 < \sigma < 1$, then $n(1-\frac{1-\sigma}{c^2}) < d < (1 - \frac{1}{\sqrt{q}-1})$ that is, a sufficient condition for the existence of such a *c*ode is:
$1-\sigma > \frac{c^2}{\sqrt{q}-1}$ (1)
but as $0 < \sigma < 1$, if $\sqrt{q} - 1 > c^2$ the code exists.
The length L_B of the inner code BS(q, r) , by proposition 1, satisfies

$$L_B \geq 8q(c + \sqrt{c-1})^2 log\frac{4q}{\varepsilon_B}$$

where $\varepsilon_B < \sigma$. By the inequality in (1) we have $q = O(c^4)$, thus $L_B = O(c^6 logc)$.
Therefore, the length of the concatenated code C = BS(q, r)oW is L = L_Bn = O(c^6log c logN).

Moreover, as the code satisfies the conditions in theorem 2 we have that $p_e \leq q^k 2^{-nD(\sigma||\varepsilon_B)}$, thus proving the theorem.

4.1 Tracing Algorithm

From the previous discussion we know that we can construct asymptotically good fingerprinting codes that allow the identification of a coalition member by a minimum Hamming distance criteria, but we have not shown how to do this identification in an efficient manner.
First note that the decoding process of the inner code, requires only $q-2$ blocks comparisons, where each block has length L = O($c^2 logc$). Therefore, the decoding time complexity for the inner code is $O(c^6 logc)$.
The decoding process for the outer code needs to recover a codeword that differs in no more than $n-n(1-\sigma)/c$ symbols from the false fingerprint. As we have seen in theorem1, we can use AG codes as outer codes, thus we can use the Guruswani-Sudan list decoding algorithm to decode them, an algorithm of *poly*(*n*) complexity.
The Guruswani-Sudan (GS) algorithm (see [7] for a detailed exposition) can be described as follows. Let C be an AG code with parameters [*n, k, d*] over IF_q.
Then, given any vector **x** = $(x_1, \ldots, x_n) \in IF_q^n$, the GS algorithm returns a list of all codewords
u = $(u_1, \ldots, u_n) \in C$ such that
$d(\mathbf{u}, \mathbf{x}) < n-\sqrt{n(n-d)}$
Thus, for our purposes, it is necessary that
$n - n(1 - \sigma)/c \leq n - \sqrt{n(n-d)}$
that is $n\sigma/c \leq n/c - \sqrt{n(n-d)}$. The last equation can be rewritten as
$\frac{1-\sigma}{c} \geq \sqrt{1-\delta}$ where $\delta = \frac{d}{n}$.
By the Tsfasman-Vladut-Zink bound, $1 - \delta \leq 1/(\sqrt{q} - 1)$, thus a sufficient condition for the existence of a code with these properties is

$$(1-\sigma)^2 > \frac{c^2}{\sqrt{q}-1}$$

but as in Theorem 1, if $\sqrt{q} - 1 > c^2$ the code exists. Therefore we have proved next theorem.

A. Theorem

Let W be an algebraic-geometric code [n, k, d] over IF_q, with N = q^k codewords, where
$d > n - n(1-\sigma)/c^2$ and $\sqrt{q} > c^2/(1-\sigma)^2 + 1$.Let V be a BS(q, r) c-secure Boneh-Shaw code with error probability $\epsilon_B < \sigma$. Then the concatenated code C = V o W is a c-secure fingerprinting code with error probability $p_e \leq q^k 2^{-nD(\sigma||\epsilon_B)}$,length L=O($c^6$log c logN) and identification algorithm complexity poly(logN).
Finally, we only want to point out that Reed-Solomon (RS) codes can be used as outer codes, obtaining reasonable lengths, however no asymptotically good codes. The RS codes are very easy to manipulate (encode/decode), but they have the not desirable property that n=q, that is, their length coincides with the cardinal of their alphabet.
Thus, when we concatenate a RS[n, k] code with a BS(n, r), that is, C=BS(n, r) ◦ RS(n, k), with have L_C= O(c^2n^2). Note that the error probability does not change.

V. Conclusion

As a conclusion we compare our construction with previous work on fingerprinting codes.
The construction presented in this co+rrespondence is a combination of the constructions of Boneh and Shaw in [5] and of Barg et al. in [8]. As one immediately sees, the resulting construction is a concatenated code where the outer code is from the same family of codes as the outer codes used in [8] and the inner

code is the code discussed in [5]. In a sense we have tried to obtain a new code by borrowing from the key features of these previous schemes.
Among the variable parameters of the construction of a fingerprinting code, that is, the number of users N, the coalition size c and the error probability ϵ, we think that c and N are in some way correlated, because the probability of great coalitions necessary increases with N, thus for asymptotic results, the behavior in c must be take in account. the goals of this paper were accomplished. The authors found a good way to use well the homomorphic property of their suggested encryption algorithm, which is simple to implement and at the same time they achieved an end-to-end encryption between the sensor nodes. They achieved bit-length reduction in the transmission process and automatically prolonged the lifetime of the WSN and saved spare bandwidth consumption.

References

[1]. Holger Karl and Andreas Willig. Protocols and Architectures for Wireless Sensor Networks. John Wiley & Sons, 2005.
[2]. Predator drones hacked in Iraq operations. http://www.cnet.com, 2009-12-24, 2009.
[3]. Gene Tsudnik Claude Castellucia, Einar Mykletun. Efficient aggregation of encrypted data in wireless sensor networks. Mobile and Ubiquitous Systems: Networking and Services, 2005. MobiQuitous 2005. The Second Annual International Conference on, 2005.
[4]. Steven Smith. Digital Signal Processing: A Practical Guide for Engineers and Scientists. Newnes, 2002.
[5]. D. Boneh and J. Shaw. Collusion-secure fingerprinting for digital data. IEEE Trans. Inform. Theory, 44(95):1897–1905, Sep. 1998.
[6]. M. Tsfasman and S. Vl˘adut. Algebraic-geometric codes. Dordrecht, The Netherlands: Kluwer, 1991.
[7]. V. Guruswami and M. Sudan. Improved decoding of Reed-Solomon codes and algebraic geometry codes. IEEE Trans. Inform. Theory, 45(6):1757–1767, Sep. 1999.
[8]. A. Barg, G. R. Blakey, and G. A. Kabatiansky. Digital fingerprinting codes: Problem statements, constructions, identification of traitors. IEEE Trans. Inform. Theory, 49(4):852–865, Apr. 2003.

