This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2009 proceedings.

A Constrained Function Based Message Authentication Scheme for Sensor Networks

Chia-Mu Yu[§†], Chun-Shien Lu[§*], and Sy-Yen Kuo[†]
[§]Institute of Information Science, Academia Sinica, Taipei, Taiwan 115, ROC
[†]Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan 106, ROC

***Abstract*—This paper presents a Constrained Function based message Authentication (CFA) scheme for wireless sensor networks, which meets all the requirements of the so-called *sensor authentication criteria*, while most of the existing schemes only achieve partial requirements. In particular, to the best of our knowledge, CFA is the first authentication scheme supporting en-route filtering with only a single packet overhead. We examine the CFA scheme from both the theoretical and experimental aspects. Our method has also been practically implemented on the TelosB compatible mote for performance evaluation.**

I. Introduction

A. Background

A Wireless Sensor Network (WSN) is composed of a large number of sensor nodes with limited resources. Since WSNs can be deployed in an unattended or hostile environment, how to design an efficient authentication scheme is of great importance to the data authenticity and integrity in WSNs. In this aspect, numerous authentication schemes have been proposed. The intuitive way to guarantee data authenticity and integrity is to use conventional public-key cryptography based digital signature scheme. Although the applicability of public-key cryptography on the WSNs has been demonstrated [4], the computation overhead is still rather high for a resource-constrained device.

Several so-called *en-route filtering schemes* have been proposed to resist the bogus event report injected by the adversary. Here, en-route filtering means that not only the destination node but also the intermediate nodes can immediately check the authenticity and integrity of the message. It is useful to mitigate Denial of Service (DoS) attack, especially path-base DoS [1], because bogus messages will be filtered out as soon as possible. For example, SEF [8], which exploits probabilistic key sharing over a partitioned key pool, and IHA [11], which verifies the going-through packet in a sophisticated hop-by-hop fashion, are proposed to authenticate the event report, respectively. Despite their effectiveness, they are not resilient to a large number of node compromises. Aiming at this problem, LBRS [9] and LEDS [6] take advantage of location information to enhance the resilience to node compromises. As to broadcast authentication, μTESLA and its variants [3], [5] can also serve message authentication well. Nevertheless, the incurred delay and vulnerability to DoS attack restrict the use of this kind of schemes.

In general, broadcast authentication is used to authenticate the message sent from the base station while en-route filtering schemes are effective only for filtering bogus event report, which is not assumed to be detected by multiple legitimate sensor nodes. Obviously, the delivery of, for example, the internal message such as control message or location information of the node itself cannot be authenticated by using broadcast authentication and en-route filtering schemes because the traffic pattern belongs to node-to-node and such contents are not known by the other nodes in nature. Thus, an authentication scheme that can efficiently authenticate the internal message is desirable. Recently, Zhang *et al*. [12] proposed an authentication scheme that can achieve such goal by exploiting so-called random perturbation. Zhang *et al*.'s scheme not only is resilient to a large number of node capture, but also supports en-route filtering.

A set of seven requirements in so-called sensor authentication criteria needed to be satisfied for an authentication scheme in WSNs are stated as follows: 1) *Resilience to a large number of node compromise* (RLNNC); 2) *Immediate authentication* (IA); 3) *Non-repudiation* (NR); 4) *Independence to network setting and deployment knowledge* (IN); 5) *Efficiency* (EFF); 6) *Scalability* (SCA); and 7) *Independence to hardware* (IH). Note that immediate authentication can be thought of as a synonym to en-route filtering. Note also that independence to hardware is important because sensor nodes maybe heterogeneous so that probably hardware setting is not able to be modified. While most of the existing schemes only satisfy some requirements in sensor authentication criteria, very recently, Zhang *et al*. [12] proposed an efficient authentication scheme based on random perturbation that satisfies the first four requirements in sensor authentication criteria. Nevertheless, the last three requirements, EFF, SCA, and IH, are not always satisfied.

B. Our Contributions

In this paper, a Constrained Function based message Authentication (CFA) Scheme for WSNs is proposed. Our contributions are twofold:

- CFA satisfies all the requirements in sensor authentication criteria, whereas most of the existing schemes satisfy only some requirments. In particular, to the best of our knowledge, CFA is the first authentication scheme supporting en-route filtering with only a single packet overhead.

[*]Contact Author: Dr. C. S. Lu (lcs@iis.sinica.edu.tw).

978-1-4244-2948-6/09/$25.00 ©2009 IEEE

- The CFA scheme is studied in both theoretical and experimental aspects. In particular, the prototype of the CFA scheme has been implemented on the TelosB compatible mote to evaluate the overhead.

II. The CFA Scheme

Since the proposed CFA scheme is constructed by making use of the pairwise key generated by the CRPV+ scheme [7], we first briefly review CRPV+ in Sec. II-B. Then, the proposed CFA scheme will be presented in the remaining sections. In this paper, nodes u, v, and i are denoted as the source node, destination node, and intermediate node, respectively.

A. System Model

Network model. We assume a WSN composed of N resource-limited sensor nodes with IDs, $\mathcal{I} \subset \mathbb{N}$. The unique ID for each node can be either arbitrarily assigned in the sensor platform, such as telosB or fixed in a specific sensing hardware when manufactured, like the MAC address on current Network Interface Cards (NICs). Although a base station or data sink is involved in data collection in a WSN, we do not rely on its trustworthiness and authenticity. In addition, arbitrary network topology is allowed in our method. The network planner, prior to sensor deployment, also cannot gain any deployment knowledge pertaining to sensors' locations.

Threat model. In this paper, sensor nodes are not equipped with tamper-resistant hardware. Thus, all the information is exposed to and can be utilized by the adversary as long as a node is captured. We also assume that the attack is mounted by the adversary immediately after sensor deployment, *i.e.*, the proposed method does not rely on the secure bootstrapping time used in [6], [10]. Two kinds of adversaries are considered in this paper, which are oblivious adversary and smart adversary. They will be described in more detail in Sec. II-G.

B. Review of the CRPV+ Scheme [7]

Let N, λ, and $\mathbb{F}_q = \{0, \ldots, q-1\}$ be the number of sensor nodes, a security parameter independent of N, a finite field, respectively. Let $A = (D \cdot G)^T$, where $D \in \mathbb{F}_q^{(\lambda+1)\times(\lambda+1)}$ is a symmetric matrix, $G \in \mathbb{F}_q^{(\lambda+1)\times N}$ is a matrix, and $(D \cdot G)^T$ is the transpose of $(D \cdot G)$. Let $K = A \cdot G$. It can be known that K must be symmetric because $A \cdot G = (D{\cdot}G)^T{\cdot}G = G^T{\cdot}D{\cdot}G = (A{\cdot}G)^T$. Before sensor deployment, proper constrained random perturbation vectors are selected and added for each row vector of A to construct a matrix W. In addition, G is selected as a Vandermonde matrix generated by a seed s. Together with s, the j-th row vector of W, $W_{j,-}$, is stored into the node j. After sensor deployment, node u can have the shared key with node v by calculating the inner product of the row vector $W_{u,-}$ and the v-th column vector $G_{-,v}$. Note that in the CRPV+ scheme G and s can be publicly known while A should be kept secret. CRPV+ can establish a pairwise key between each pair of sensor nodes without needing any communication. This property is an essential part in constructing the proposed CFA scheme, because establishing a key via communications incurs authentication problem, leading to circular dependency. CRPV+ is also resilient to a large number of node compromises so that the complexity for breaking the CRPV+ scheme is $\Omega(2^{\lambda+1})$.

C. Basic Idea

In the CFA scheme, the network planner, before sensor deployment, selects a secret polynomial $f(x,y,z,w)$ from the set $\mathfrak{F}$ (defined in Eq. (1)), whose coefficients constitute the security basis of CFA. For simplicity, we assume that the degree of each variable in $f(x,y,z,w)$ is the same, which is d. For each node u, the network planner constructs two polynomials, $f_{u,1}(y,z,w) = f(u,y,z,w)$ and $f_{u,2}(x,z,w) = f(x,u,z,w)$. Recall that the coefficients of $f(x,y,z,w)$ should be kept as secret. Since directly storing these two polynomials enables the adversary to obtain the coefficients of $f(x,y,z,w)$ by capturing a few nodes, thus, the authentication polynomial $auth_u(y,z,w)$ and verification polynomial $verf_u(x,z,w)$ should be constructed from the polynomials $f_{u,1}(y,z,w)$ and $f_{u,2}(x,z,w)$, respectively, by adding independent perturbation polynomials. Afterwards, the authentication and verification polynomials, instead of $f_{u,1}(y,z,w)$ and $f_{u,2}(x,z,w)$, are stored in node u. For source node u, the *message authentication code* attached to the message m is calculated according to its own authentication polynomial. Let *verification number* be the result calculated from the verification polynomial $verf_u(x,z,w)$ by applying the claimed source node ID and the shared pairwise key to the claimed source node and the hashed message, respectively. The received node considers the received message authentic and intact if and only if the *verification difference*, which is the difference between the message authentication code and its calculated verification number, is within a certain predetermined range.

The techniques used in the CFA scheme are different from the ones used in [12], except the fact that both rely on polynomial evaluation. In [12], due to the improper use of perturbation, the nodes' IDs should be forced to be changed, resulting in the limitation of hardware dependence. On the other hand, because arbitrary secret polynomial can be used in [12], immediate authentication can be achieved only if the message authentication code forms a polynomial. In the CFA scheme, the secret polynomial $f(x,y,z,w)$ is selected such that certain properties are satisfied. By using such a kind of secret polynomials, the message authentication code can be reduced from a polynomial size to a single number, resulting in less communication overhead (packet overhead). In addition, whereas the pairwise key has been considered useless in providing either immediate authentication or resilience to node compromises in previous methods, in this paper we find that the pairwise key is helpful in enhancing the security while retaining the property of immediate authentication.

D. Off-line Step of CFA scheme

Before deploying sensor nodes, the network planner picks a parameter q from which a finite field $\mathbb{F}_q$ is built. Let $\mathcal{I}$ be the set of node IDs. Let $\ell, \ell_{\mathcal{I}}, \ell_k$, and ℓ_h, respectively, represent the least number of bits sufficient to represent q, node ID, pairwise key, and hash value. In addition, a security parameter $r < \ell$ is also selected. Then, the secret polynomial $f(x, y, z, w)$ is selected from the *constrained function set*, $\mathfrak{F}$, where

$$\begin{aligned}\mathfrak{F} = \{&f(x, y, z, w) \| f(x, y, z, w) - f(x, y', z', w)| \leq 2^{r-1},\\ &|f(x, y, z, w) - f(x', y', z', w')| \geq 3 \cdot 2^{r-1} - 1,\\ &0 \leq f(x, y, z, w) \leq 2^r, x, y \in \mathcal{I}, 0 \leq z \leq 2^{\ell_k} - 1,\\ &0 \leq w, w' \leq 2^{\ell_h} - 1, x' \neq x, y' \neq y, z' \neq z, r < \ell\}.\end{aligned} \tag{1}$$

The authentication polynomial, $auth_u(y, z, w) = f(u, y, z, w) + n_{u,\mathfrak{a}}(z)$, and verification polynomial, $verf_u(x, z, w) = f(x, u, z, w) + n_{u,\mathfrak{v}}(z)$, are stored in each node u, where polynomials $n_{u,\mathfrak{a}}(z)$ and $n_{u,\mathfrak{v}}(z)$, used for perturbation, are randomly selected from the *authentication perturbation set*, $\mathfrak{N}_\mathfrak{a} = \{n(z) | 0 \leq n(z) \leq 2^{r-2} - 1, 0 \leq z \leq 2^{\ell_k} - 1\}$, and the *verification perturbation set*, $\mathfrak{N}_\mathfrak{v} = \{n(z) | 0 \leq n(z) \leq 2^{r-1} - 1, 0 \leq z \leq 2^{\ell_k} - 1\}$, respectively. Though the sets $\mathfrak{F}$, $\mathfrak{N}_\mathfrak{a}$, and $\mathfrak{N}_\mathfrak{v}$ appear to be artificial, they guarantee the efficiency and feasibility of immediate authentication of CFA. In addition, the construction of these three sets may be time- and energy-consuming. Nevertheless, it can be acceptable because this construction is performed only by the network planner, instead of sensor nodes. An efficient method to construct the polynomials in $\mathfrak{F}$ will be shown in Sec. II-F. The off-line procedure of CFA is described in Fig. 1.

```
Algorithm: CFA-Off-line-Step($q$, $r$)
Randomly picks a secret polynomial $f(x, y, z, w) \in \mathfrak{F}$
for each node $u$
    Randomly picks $n_{u,\mathfrak{a}}(z) \in \mathfrak{N}_\mathfrak{a}$ and $n_{u,\mathfrak{v}}(z) \in \mathfrak{N}_\mathfrak{v}$
    Store $auth_u(y, z, w) = f(u, y, z, w) + n_{u,\mathfrak{a}}(z)$
    Store $verf_u(x, z, w) = f(x, u, z, w) + n_{u,\mathfrak{v}}(z)$
```

Fig. 1. Off-line Step of CFA.

E. On-line Step of CFA scheme

After sensor deployment, the sensor node may work as a source node, intermediate node, or destination node depending on whether the message is to be sent or verified. In the following, we describe the operations one should perform when the node acts as different roles. It should be noted that the pairwise key $K_{u,v} = K_{v,u}$, used here, is constructed by applying the CRPV+ scheme [7] on nodes u and v, respectively.

Source node (Message transmission). When node u wants to send a message m to node v, it calculates the message authentication code

$$MAC_u(v, m) = auth_u(v, K_{u,v}, h(m)) + n_{u,s},$$

where $n_{u,s}$ is randomly picked from the set $\{0, \ldots, 2^{r-2}\}$. Then, the packet $\mathcal{M} = \langle u, v, m, MAC_u(v, m) \rangle$ is sent out. Note that the message authentication code $MAC_u(v, m)$ is only a number here.

Destination node (Message verification). After receiving the packet $\mathcal{M}$, the destination node v first calculates the verification number:

$$verf_v(u, K_{u,v}, h(m)),$$

according to its own verification polynomial $verf_v(x, z, w)$ and then calculates the verification difference, VD:

$$VD = |verf_v(u, K_{u,v}, h(m)) - MAC_u(v, m)|.$$

If VD is within the range $\{0, \ldots, 2^{r-1} - 1\}$, then the authenticity and integrity of the packet $\mathcal{M}$ is successfully verified. Otherwise, the packet $\mathcal{M}$ is dropped. The principle behind this step is as follows:

$$\begin{aligned}&verf_v(u, K_{v,u}, h(m)) - MAC_u(v, m)\\ =&(f(u, v, K_{v,u}, h(m)) + n_{v,\mathfrak{v}}(K_{v,u}))\\ &- (f(u, v, K_{u,v}, h(m)) + n_{u,\mathfrak{a}}(K_{u,v}) + n_{u,s})\\ =&(f(u, v, K_{v,u}, h(m)) - f(u, v, K_{u,v}, h(m)))\\ &+ (n_{i,\mathfrak{v}}(K_{v,u}) - (n_{u,\mathfrak{a}}(K_{u,v}) + n_{u,s})\\ =&n_{i,\mathfrak{v}}(K_{v,u}) - (n_{u,\mathfrak{a}}(K_{u,v}) + n_{u,s}).\end{aligned} \tag{2}$$

From the rules of constructing authentication and verification polynomials, we know that $n_{i,\mathfrak{v}}(K_{i,u}) \in \{0, \ldots, 2^{r-1} - 1\}$, $n_{u,\mathfrak{a}}(K_{u,v}) \in \{0, \ldots, 2^{r-2} - 1\}$, and $n_{u,s} \in \{0, \ldots, 2^{r-2}\}$. Thus, when $\mathcal{M}$ is genuine, the verification difference $VD = |verf_v(u, K_{u,v}, h(m)) - MAC_u(v, m)|$ must be within $\{0, \ldots, 2^{r-1} - 1\}$.

Intermediate node (Message verification). After receiving the packet $\mathcal{M}$, the intermediate node i first calculates $verf_i(u, K_{i,u}, h(m))$ according to its own verification polynomial $verf_i(x, z, w)$ and then calculates the verification difference $VD = |verf_i(u, K_{i,u}, h(m)) - MAC_u(v, m)|$. If VD is within the range $\{0, \ldots, 2^r - 1\}$, then the authenticity and integrity of the packet $\mathcal{M}$ is successfully verified, and the packet $\mathcal{M}$ will be forwarded by node i. Otherwise, the packet $\mathcal{M}$ is dropped. The principle behind this step is as follows. When genuine packet $\mathcal{M}$ is received, we can get

$$\begin{aligned}&verf_i(u, K_{i,u}, h(m)) - MAC_u(v, m)\\ =&(f(u, i, K_{i,u}, h(m)) + n_{i,\mathfrak{v}}(K_{i,u}))\\ &- (f(u, v, K_{u,v}, h(m)) + n_{u,\mathfrak{a}}(K_{u,v}) + n_{u,s})\\ =&(f(u, i, K_{i,u}, h(m)) - f(u, v, K_{u,v}, h(m)))\\ &+ (n_{i,\mathfrak{v}}(K_{i,u}) - (n_{u,\mathfrak{a}}(K_{u,v}) + n_{u,s}).\end{aligned} \tag{3}$$

By the construction of $\mathfrak{F}$, we know

$$|f(u, i, K_{i,u}, h(m)) - f(u, v, K_{u,v}, h(m))| \leq 2^{r-1}. \tag{4}$$

In addition, from the rules of constructing authentication and verification polynomials, we know that $n_{i,\mathfrak{v}}(K_{i,u}) \in \{0, \ldots, 2^{r-1} - 1\}$, $n_{u,\mathfrak{a}}(K_{u,v}) \in \{0, \ldots, 2^{r-2} - 1\}$, and $n_{u,s} \in \{0, \ldots, 2^{r-2}\}$. Therefore, the verification difference VD must be within $\{0, \ldots, 2^r - 1\}$.

On the other hand, consider the modified packet,

$$\mathcal{M}' = \langle u', v, m, MAC_u(u'', m) \rangle, \tag{5}$$

where u' means a node ID the adversary pretends to be, and u'' can be arbitrary. Note that we only consider the adversary who exploits the information obtained from a single captured node u, and focus on the use of constructed set $\mathfrak{F}$. The verification procedure is as follows:

$$\begin{aligned} & verf_i(u', K_{i,u'}, h(m)) - MAC_u(u'', m) \\ =& (f(u', i, K_{i,u'}, h(m)) + n_{i,\mathfrak{v}}(K_{i,u'})) \\ & - (f(u, u'', K_{u',u''}, h(m)) + n_{u,\mathfrak{a}}(K_{u',u''}) + n_{u,s}) \\ =& (f(u', i, K_{i,u'}, h(m)) - f(u, u'', K_{u',u''}, h(m))) \\ & + (n_{i,\mathfrak{v}}(K_{i,u'}) - (n_{u,\mathfrak{a}}(K_{u',u''}) + n_{u,s}). \end{aligned} \tag{6}$$

By the construction of $\mathfrak{F}$, we know

$$|f(u', i, K_{i,u'}, h(m)) - f(u, u'', K_{u',u''}, h(m))| \geq 3 \cdot 2^{r-1} - 1. \tag{7}$$

In addition, from the construction of authentication and verification polynomials, we know that $n_{i,\mathfrak{v}}(K_{i,u'}) \in \{0, \ldots, 2^{r-1} - 1\}$, $n_{u,\mathfrak{a}}(K_{u',u''}) \in \{0, \ldots, 2^{r-2} - 1\}$, and $n_{u,s} \in \{0, \ldots, 2^{r-2}\}$. Therefore, the verification difference VD must be not within $\{0, \ldots, 2^r - 1\}$ and the packet $\mathcal{M}'$ will be dropped. The on-line procedure of CFA is described in Fig. 2.

```
Algorithm: CFA-On-line-Step
Scenario: node $u$ sends a message $m$ to node $v$
Source node $u$:
  1. Calculate $K_{u,v}$ and $h(m)$
  2. Compute $MAC_u(v, m) = auth_u(v, K_{u,v}, h(m)) + n_{u,s}$,
     where $n_{u,s}$ is randomly picked from $\{0, \ldots, 2^{r-2} - 1\}$
  3. Send the packet $\mathcal{M} = \langle u, v, m, MAC_u(v, m) \rangle$
Intermediate node $i$ (on receiving $\mathcal{M}$):
  1. Calculate $K_{u,v}$ and $h(m)$
  2. Calculate $x = |verf_v(u, K_{i,u}, h(m)) - MAC_u(v, m)|$
  3. if $x \in \{0, \ldots, 2^r - 1\}$
     then forwarding $\mathcal{M}$ else drop $\mathcal{M}$
Destination node $v$ (on receiving $\mathcal{M}$):
  1. Calculate $K_{u,v}$ and $h(m)$
  2. Calculate $x = |verf_v(u, K_{v,u}, h(m)) - MAC_u(v, m)|$
  3. if $x \in \{0, \ldots, 2^{r-1} - 1\}$
     then accept $\mathcal{M}$ else drop $\mathcal{M}$
```

Fig. 2. On-line Step of CFA.

Example 1. *Consider a network with* 256 *nodes whose IDs are in the set* $\{1, \ldots, 256\}$. *Let* $q = 2^{32}$, $r = 19$, *and* $0 \leq \ell_k, \ell_h \leq 8$. *Select* $f(x, y, z, w) = 262144x^2 + y^2 + z^2 + w^2$ *as the secret polynomial. Suppose that node* 1 ($u = 1$) *wants to transmit a message* m *to node* 5 *through node* 3*, and the hash* $h(m)$ *of the message* m *is* 12*. Assume that the authentication polynomial stored in node* 1 *is* $auth_1(y, z, w) = y^2 + z^2 + w^2 + 262144 + n_{1,\mathfrak{a}}$*, where* $n_{1,\mathfrak{a}} = -z^2 + 256z$*. In addition, assume that the verification polynomial stored in node* 3 *is* $verf_3(x, z, w) = 262144x^2 + z^2 + w^2 + 9 + n_{3,\mathfrak{v}}$*, where* $n_{3,\mathfrak{v}} = 3z^2 - 356z + 10600$*, and the verification polynomial stored in node* 5 *is* $verf_3(x, z, w) = 262144x^2 + z^2 + w^2 + 25 + n_{5,\mathfrak{v}}$*, where* $n_{5,\mathfrak{v}} = 4z^2 - 570z + 20400$*. Let* $K_{1,3} = K_{3,1} = 56$ *and* $K_{1,5} = K_{5,1} = 120$*. Under these circumstances, by choosing* $n_{1,s} = 4$*, node* 1 *calculates* $MAC_1(5, m) = 293037$ *and sends* $\langle 1, 5, 12, 293037 \rangle$*. Node* 3 *checks the authenticity and integrity of the message* m *by calculating* $verf_3(1, 56, 12) = 265505$ *and obtains* $VD = |265505 - 293037| = 27532$*. Since* $VD \in \{0, \ldots, 2^{18} - 1\}$*, node* 3 *forwards the packet* $\langle 1, 5, 12, 293037 \rangle$*. On receiving* $\langle 1, 5, 12, 293037 \rangle$*, node* 5 *computes* $verf_5(1, 120, 12) = 286313$ *and obtain* $VD = |286313 - 293037| = 6724$*. Because* VD *is within* $\{0, \ldots, 2^{18} - 1\}$*, the authenticity and integrity of the message* m *is successfully verified.*

F. Constructing the Constrained Function Set

The effectiveness and efficiency of the proposed CFA scheme rely on the constrained function set $\mathfrak{F}$, the authentication perturbation set $\mathfrak{N}_{\mathfrak{a}}$, and the verification perturbation set $\mathfrak{N}_{\mathfrak{v}}$. As the construction of $\mathfrak{N}_{\mathfrak{a}}$ and $\mathfrak{N}_{\mathfrak{v}}$ is relatively easy, in this section, we focus on constructing $\mathfrak{F}$.

A naive algorithm for constructing the constrained function set $\mathfrak{F}$ is the exhaustive search. When the coefficients of the polynomials belonging to $\mathfrak{F}$ are constrained with a finite field $\mathbb{F}_q$, the number of all possible four-variate t-degree polynomials is $q^{(t+1)^4}$. Thus, $O(q^{2 \cdot (t+1)^4})$ tests are required because there are $q^{(t+1)^4}$ four-variate t-degree polynomials, each of which needs the check whether the other $q^{(t+1)^4} - 1$ polynomials satisfy the constraint $|f(x, y, z, w) - f(x', y', z', w')| \geq 3 \cdot 2^{r-1} - 1$ in $\mathfrak{F}$. The construction of $\mathfrak{F}$ will be accomplished by the network planner that is usually assumed to be resource-abundant before sensor deployment. However, exhaustive search is a feasible but not sufficiently efficient method. We observe that, in some cases, a restricted polynomial is sufficient for our use, and the search for the restricted polynomial can accelerate the construction of $\mathfrak{F}$. Hence, in the following, we emphasize how to efficiently construct a set of restricted polynomials, $\mathfrak{F}'$.

Let $\mathfrak{F}'$ be the *weak constrained function set* as follows:

$$\begin{aligned} \mathfrak{F}' = \{ & f(x, y, z, w) | |f(x, y, z, w) - f(x, y', z', w)| \leq 2^{r-1}, \\ & |f(x, y, z, w) - f(x', y', z', w')| \geq 3 \cdot 2^{r-1} - 1, \\ & 0 \leq f(x, y, z, w) \leq 2^r, \\ & f(x, y, z, w) = \alpha_1 x^d + \alpha_2 y^d + \alpha_3 z^d + \alpha_4 w^d, \\ & \alpha_1, \alpha_2, \alpha_3, \alpha_4 \in \mathbb{Z}, d \in \mathbb{N}, x, y \in \mathcal{I}, 0 \leq z \leq 2^{\ell_k} - 1, \\ & 0 \leq w, w' \leq 2^{\ell_h} - 1, x' \neq x, y' \neq y, z' \neq z, r < \ell \}. \end{aligned} \tag{8}$$

Obviously, the weak constrained function set is a subset of the constrained function set, *i.e.*, $\mathfrak{F}' \subset \mathfrak{F}$, because one additional constraint $f(x, y, z, w) = \alpha_1 x^d + \alpha_2 y^d + \alpha_3 z^d + \alpha_4 w^d$ is added into $\mathfrak{F}$. The consideration of $\mathfrak{F}'$ has an easier construction in the sense that the search space for the polynomials $f(x, y, z, w)$ is reduced. Since in most cases, the consideration of $\mathfrak{F}'$ is sufficient to guarantee the security, we focus on the construction of $\mathfrak{F}'$. Here, a lemma that describes the construction of $\mathfrak{F}'$ is given in the following.

Lemma 1. *If the coefficients α_1, α_2, α_3, and $\alpha_4 \in \mathbb{F}_q$ are selected such that the following three constraints*

$$\alpha_1 x^d + \alpha_4 w^d \geq 2^r - 1, \tag{9}$$

$$\alpha_2 y^d + \alpha_3 z^d \leq -2^r + 1, \tag{10}$$

and

$$\min_{x,w}\left(\alpha_1 x^d + \alpha_4 w^d\right) > \min_{y,z}\left(\alpha_2 y^d + \alpha_3 z^d\right) \tag{11}$$

can be satisfied, then the polynomial $f(x, y, z, w) = \alpha_1 x^d + \alpha_2 x^d + \alpha_3 x^d + \alpha_4 x^d$ belongs to the weak constrained function set, $\mathfrak{F}'$.

Proof: From the definition of $\mathfrak{F}'$, shown in Eq. (8), it can be known that the polynomial $f(x, y, z, w)$ belonging to $\mathfrak{F}'$ must satisfy with the constraint

$$0 \leq f(x, y, z, w) \leq 2^r. \tag{12}$$

Based on the constraint $|f(x, y, z, w) - f(x, y', z', w)| \leq 2^{r-1}$ shown in Eq. (8), we have:

$$\begin{aligned}
&\alpha_1 x^d + \alpha_2 y^d + \alpha_3 z^d + \alpha_4 w^d \\
&-\alpha_1 x^d - \alpha_2 (y')^d - \alpha_3 (z')^d - \alpha_4 w^d \\
&=\alpha_2 (y^d - (y')^d) + \alpha_3 (z^d - (z')^d) \leq 2^{r-1}.
\end{aligned} \tag{13}$$

Here, we consider $f(x, y, z, w) - f(x, y', z', w) \leq 2^{r-1}$, instead of $|f(x, y, z, w) - f(x, y', z', w)| \leq 2^{r-1}$, since it can trivially implies $|f(x, y, z, w) - f(x, y', z', w)| \leq 2^{r-1}$. In addition, based on the constraint $|f(x, y, z, w) - f(x', y', z', w')| \geq 3 \cdot 2^{r-1} - 1$ shown in Eq. (8), we have:

$$\begin{aligned}
&\alpha_1 (x^d - (x')^d) + \alpha_2 (y^d - (y')^d) \\
&+\alpha_3 (z^d - (z')^d) + \alpha_4 (w^d - (w')^d) \geq 3 \cdot 2^{r-1} - 1.
\end{aligned} \tag{14}$$

The reason to the ignorance of absolute operation here is the same as in Eq. (14). Subtracting Eq. (13) from Eq. (14), we can obtain:

$$\alpha_1 (x^d - (x')^d) + \alpha_4 (w^d - (w')^d) \geq 2^r - 1. \tag{15}$$

Both conditions $\alpha_1 x^d \geq \alpha_1 (x^d - (x')^d)$ and $\alpha_4 w^d \geq \alpha_4 (w^d - (w')^d)$ always hold based on the limitation of x and w in Eq. (8). Thus, we have the following relation from Eq. (15):

$$\alpha_1 x^d + \alpha_4 w^d \geq 2^r - 1. \tag{16}$$

Hence, subtracting Eq. (16) from the second inequality of Eq. (12), we obtain:

$$\alpha_2 y^d + \alpha_3 z^d \leq -2^r + 1 + 2^r = 1. \tag{17}$$

Since the first inequality of Eq. (13) should be satisfied, we know that an additional constraint must be considered:

$$\min_{x,w}\left(\alpha_1 x^d + \alpha_4 w^d\right) > \min_{y,z}\left(\alpha_2 y^d + \alpha_3 z^d\right). \tag{18}$$

Otherwise, the polynomial constructed by using Eq. (16) and Eq. (17) can satisfy the first inequality of Eq. (13). As a result, the lemma is proven by considering Eqs. (16), (17), and (18). □

TABLE I
IMPLEMENTATION RESULTS OF CFA

Scheme	Time	Cycle Count	Flash Memory	RAM
CFA	2.839ms	22713	408B	1220B

Example 2. *Assume that $d = 2$, $q = 2^{32} - 1$, $r = 19$, $1 \leq x, y \leq 256$, and $0 \leq z, w \leq 255$. From Lemma 1, it can be known that the polynomial $\alpha_1 x^2 + \alpha_2 y^2 + \alpha_3 z^2 + \alpha_4 w^2$, whose coefficients α_1, α_2, α_3, and α_4 satisfy Eq. (9), Eq. (10), and Eq. (11), belongs to $\mathfrak{F}'$. One of settings for the coefficients $\alpha_1 \sim \alpha_4$ can be calculated as $\alpha_1 = 2^{28} - 1$, $\alpha_2 = -2^{19} - 1$, and $\alpha_3 = \alpha_4 = 0$.*

G. Performance Evaluation

The prototype of the proposed CFA scheme was implemented on the Taroko[1] mote (a TelosB[2] compatible mote made in Taiwan). Note that the programming tool we used was native C compiler on IAR Embedded Workbench[3] 3.40.1.9, instead of TinyOS. In our experiments, the parameters used were the same as the ones described in Example 2, except certain implementation details. We used the diagnostic and profiling outputted from IAR Embedded Workbench to estimate storage and computation overhead. The implementation results are shown in Table. I. It should be noted that the implementation results are preliminary and are not optimized. Complete implementation of the CFA scheme will be presented in the further version.

Storage Overhead. Since CFA exploits the pairwise key constructed by using the CRPV+ scheme [7], the storage overhead comprises the memory usage incurred by the authentication and verification procedures, and CRPV+ scheme. In the authentication and verification procedures, two trivariate polynomials are needed to be stored, while in the CRPV+ scheme a $(\lambda + 1)$-dimensional vector is needed to be stored. Therefore, constant size storage overhead, $O(\lambda + d^3)$, is required in CFA.

Computation Overhead. For a source node, the computation overhead comes from the calculation of pairwise key and message authentication code. The former involves univariate polynomial evaluation, while the latter involves trivariate polynomial evaluation. Even if a complicated algorithm is not used, a naive method, in general, can accomplish evaluation by using $O(d + 3d^3)$ multiplications and $O(d^3)$ additions, resulting in constant computation overhead. On the other hand, the computation overhead for intermediate nodes and destination node is the same as the one for the source node, except that one additional substraction is required.

Communication Overhead (Packet Overhead). In CFA, it does not incur additional communication when an authenticated message is required to be sent from node u to node v. Thus, when communication overhead of CFA scheme is

[1] Micro-Controller: TI MSP430F1611, Flash Memory: 48KB+256B, RAM: 10KB, Radio Chipset: ChipCon CC2420.

[2] Availiable at: www.xbow.com/Products/Product_pdf_files/Wireless_pdf/TelosB_Datasheet.pdf

[3] Availiable at: www.iar.com

mentioned, we focus on the additionally increased message length resulted from the use of CFA. Since the additional item attached to the message m is only a number, *i.e.*, $MAC_u(v, m)$, the communication overhead is a constant, which is optimal.

Security. We consider *oblivious adversary* and *smart adversary* in this paper. The oblivious adversary tries to deceive a sensor node v that the received message is sent from another legitimate sensor node u' without modification. The smart adversary can capture a subset of sensor nodes, $\mathcal{C} = \{c_1, \ldots, c_{|\mathcal{C}|} | c_i \in \mathcal{I}, i \neq j, c_i \neq c_j, 1 \leq i, j \leq |\mathcal{C}|\}$, and can eavesdrop on the transmitted messages throughout the network. The goal of smart adversary is to deceive a sensor node v that the received message is sent from another legitimate sensor node $u' \notin \mathcal{C}$ without modification. The ability of the CFA scheme in resisting both oblivious adversary and smart adversary are described in Lemma 2 and Lemma 3, respectively.

Lemma 2. *If the packet $\mathcal{M} = \langle u, v, m, MAC_u(v, m)\rangle$ is modified by an oblivious adversary to $\mathcal{M}'$, where either u is replaced by u' or m is replaced by m', then the probability that the packet $\mathcal{M}'$ is, respectively, verified by the destination node and the intermediate node as valid is $O(2^{r-\ell-d\cdot(\ell_{\mathcal{I}}+\ell_k+\ell_h)})$ and $O(2^{r-\ell-d\cdot(\ell_{\mathcal{I}}+\ell_k+\ell_h)+1})$.*

Lemma 3. *The complexity that the packet $\mathcal{M} = \langle u, v, m, MAC_u(v, m)\rangle$ is modified by a smart adversary to $\mathcal{M}'$, where either u is replaced by u' or m is replaced by m' and the packet $\mathcal{M}'$ is verified as valid by the destination node v is $\Omega(2^{r\cdot(\lambda+1)\cdot(d+1)})$.*

Comparison. The performance comparison is depicted in Table. II. When focusing on the comparison between Zhang *et al.*'s scheme [12], which is considered to be the best in the literature, and the proposed CFA scheme, we can see that three requirements in the sensor authentication criteria can be improved in our method. In Zhang *et al.*'s scheme, the node ID is artificial, leading to hardware dependency and therefore limited applications. Moreover, due to the artificiality of node ID, when the network size is increased, the number of bits needed to represent the node ID is quickly increased as well, resulting in increased communication overhead and limited scalability. Finally, it is worthy of mentioning that the proposed CFA scheme achieves constant communication overhead. In particular, only one single number representing message authentication code is sufficient, while the packet size in [12], depending on security level, will be lengthy. It should be noted that message length is important for some WSN settings with constraint on packet size. For example, in TinyOS the default packet size is 36 bytes. Hence, increased message length induces more packets required, resulting in higher energy consumption. Moreover, as indicated in [13], once the message can be constructed only if multiple packets are received, the adversary can launch DoS attack to overrun the buffer of the sensor node by sending a large number of false packets. In CFA, since message authentication code can always encapsulated into a single packet, such vulnerability can be avoid, whereas the other existing schemes could have such vulnerability.

TABLE II
COMPARISONS BETWEEN DIFFERENT AUTHENTICATION SCHEMES

	RLNNC	IA	NR	IN	EFF	SCA	IH
SEF [8]	△	✓	×	×	△	✓	✓
IHA [11]	△	✓	×	△	△	✓	✓
LBRS [9]	✓	✓	×	×	△	✓	✓
LEDS [6]	✓	✓	×	×	△	✓	✓
Zhang *et al.* [12]	✓	✓	✓	✓	△	△	×
CFA (this paper)	✓	✓	✓	✓	✓	✓	✓

III. CONCLUSION

The contribution of this paper is to propose a Constrained Function based message Authentication (CFA) scheme to satisfy all the requirements in so-called sensor authentication criteria. Some theoretical analyses and preliminary implementation results are given to demonstrate the efficiency and effectiveness of CFA.

Acknowledgment: Chia-Mu Yu and Chun-Shien Lu were supported by NSC 97-2221-E-001-008. Sy-Yen Kuo was supported by NSC 96-2628-E-002-138-MY3.

REFERENCES

[1] J. Deng, R. Han, S. Mishra. Defending against Path-based DoS Attacks in Wireless Sensor Networks. In *ACM SASN*, 2005.

[2] D. Liu, P. Ning, "Establishing Pairwise Keys in Distributed Sensor Networks," *ACM CCS*, 2003.

[3] D. Liu and P. Ning, "Multi-Level μTESLA: Broadcast Authentication for Distributed Sensor Networks," *ACM Transactions in Embedded Computing Systems (TECS)*, Vol. 3, No. 4, pages 800-836, November 2004.

[4] D. J. Malan, M. Walsh, and M. D. Smith, "Implementaing Public-Key Infrastructure for Sensor Networks," *ACM Trans. on Sensor Networks*, 2008.

[5] A. Perrig, R. Szewczyk, V. Wen, D. Culler, and Doug Tygar, "SPINS: Security Protocols for Sensor Networks," *ACM MobiCom*, 2001.

[6] K. Ren, W. Lou, and Y. Zhang, "LEDS: providing location-aware end-to-end data security in wireless sensor networks," *IEEE INFOCOM*, 2006.

[7] C. M. Yu, T. Y. Chi, C. S. Lu, and S. Y. Kuo, "A Constrained Random Perturbation Vector-Based Pairwise Key Establishment Scheme for Wireless Sensor Networks," *ACM MobiHoc*, 2008.

[8] F. Ye, H. Luo, S. Lu, and L. Zhang, "Statistical En-route Filtering of Injected False Data in Sensor Networks," *IEEE INFOCOM*, 2004.

[9] H. Yang, F. Ye, Y. Yuam, S. Lu, and W. Arbaugh, "Toward Resilient Security in Wireless Sensor Networks," *ACM MobiHoc*, 2005.

[10] S. Zhu, S.Setia, and S.Jajodia. LEAP: Efficient Security Mechanisms for Large-Scale Distributed Sensor Networks. In *ACM CCS*, 2003.

[11] S. Zhu, S. Setia, S. Jajodia, and P. Ning, "An Interleaved Hop-by-Hop Authentication Scheme for Filtering False Data in Sensor Networks," *IEEE S&P*, 2004.

[12] W. Zhang, N. Subramanian, and G. Wang, "Lightweight and Compromise-Resilient Message Authentication in Sensor Networks," *IEEE INFOCOM*, 2008.

[13] W. Zhang, H. Song, S. Zhu, and G. Cao, "Least Privilege and Privilege Deprivation: Towards Tolerating Mobile Sink Compromises in Wireless Sensor Networks," *ACM MobiHoc*, 2005.