EAI Endorsed Transactions
on Mobile Communications and Applications
Research Article EAI.EU

E-HMAC: An Efficient Secure Homomorphic MAC Scheme for NC-Enabled WSNs

Haythem Hayouni*

SupCom, University of Carthage, Tunisia

Abstract

The main goal of Network Coding (NC) is to find an optimal transmission of data in a network. NC presents an advantage for wireless sensor networks (WSNs) in term of network lifetime. However, the Network Coding-enabled WSNs are affected by various attacks, such as pollution attacks. Many HMAC schemes have been proposed in the literature to secure packets against pollution attacks. In 2015, Esfahani et al. proposed a dual-homomorphic MAC scheme based to the construction of two different MACs to ensure the integrity of coded packets. Their solution has many weaknesses in terms of security against tag pollution attacks. In this paper, we improve their scheme by proposing a novel HMAC scheme for NC enabled WSNs, called E-HMAC, based on multi-linear space to check the integrity of coded packets. The simulation results demonstrate the ability of our proposed scheme to secure the coded packets with a low key storage overhead and communication overhead, compared to Esfahani et al.'s scheme.

Keywords: Network Coding, Wireless sensor networks, Homomorphic MAC, linear mapping, pollution attacks

Received on 06 May 2021, accepted on 26 September 2021, published on 29 September 2021

doi: 10.4108/eai.29-9-2021.171189

1. Introduction

WSNs [1] consist of a set of devices having limited computing resources. This type of network has attracted much attention in recent years, not only in academia but also in industry, for the study and development of a number of potential applications. However, the resource constraint is the most important feature of this network. Network Coding (NC) find an optimal transmission of data in a network. Network Coding can also improve network resiliency against attacks. Wireless Sensor Networks (WSNs) can benefit from the benefits of NC.

1.1 Motivations

In many applications of Network Coding-enabled WSNs, data can be threatened by external events that should not occur during normal network operation. Among these attacks, we find pollution attacks [2]. There are two types of pollution attacks: data pollution attack and tag pollution attack. In data pollution attack, the mission of an adversary is to insert fake data and to realize the verification of other innocent sensors which causes. In tag pollution attack, the objective of adversary is to get correct data packets be beaked as false and be isolated by intermediate sinks or nodes, which discard the correct packets. If a data pollution attack is not detected at the forwarders nodes, the base station cannot be able to verify if the received message is correct or not, and cannot check the source messages correctly. In WSN, as long as the polluted packets propagates via recording, a small number of these packets can affect the security of large number of downstream nodes. There are several cryptographic methods [3] providing the security and integrity of transmitted data such as Homomorphic MAC [4][5]. In 2015, Esfahani et al. proposed a dual homomorphic MAC scheme based the generation of two different MACs to ensure the integrity of

*Email: haythem.hayouni@supcom.tn

coded packets. Their solution has many drawbacks in terms of security against tag pollution attacks. Among the proposed schemes in the literature, to prevent data against pollution attacks, Homomorphic MAC is considered as a low-complexity method for this. But, the generated MACs can be polluted during their transmission between nodes until the base station, which make the solution of HMAC vulnerable to tag pollution attacks. For this reason, it is necessary to produce a secure homomorphic MACs which mitigate both partially tag pollution attacks and data pollution attacks. In this paper, we improve the Esfahani et al.'s scheme [9], which is the most well-known tag pollution scheme taking advantage of homomorphic MACs, by proposing for Network Coding-enabled WSNs an efficient HMAC scheme, called E-HMAC, to ensure the integrity of packets transmitted in the network. The objective is to protect data against pollution attacks.

1.2. Our contributions

The main contributions of our proposal are as follows:

- We review the state-of-the-art regarding the multiple proposals that use homomorphic MAC to secure the transmitted messages against pollution attacks in NC-enabled networks.
- We reviewed the scheme of Esfahani et al. and pointed out the weaknesses in terms of security against tag pollution attacks.
- We propose an efficient HMAC scheme, called E-HMAC, to ensure the integrity of packets transmitted in the network, by we improving the Esfahani et al.'s scheme.
- We introduce some new concepts to protect data against pollution attacks.
- We secure the coded packets with a low consumption of network resources.
- We prove our scheme secure using a security analysis against Data Pollution Attack and Tag Pollution Attack by proving some theorems.
- We evaluate the performance of the proposed scheme with respect to its key storage and communication overhead.
- We compare our findings to Esfahani et al.'s scheme and discuss our results.

The rest of the paper is outlined as follows. In Section 2, we presented some related works for ensuring security against pollution attacks, based on HMAC method. In Section 3, some preliminaries are presented and discussed. Section 4 reviews Esfahani et al.'s scheme and then Section 5 proves the security shortcomings of Esfahani et al.'s scheme. Section 6 presents the proposed scheme E-HMAC. Section 7 evaluates the security analysis of E-HMAC. Section 8 presents the results of the performance analysis of the E-HMAC compared with Esfahani et al.'s scheme. Finally, Section 9 concludes the paper.

2. Related works

However, jointly ensuring the security, such as integrity, of the packets exchanged in WSNs is a challenge as the nodes are deployed in hostile environments. Sensor nodes are exposed to several attacks such as pollution attacks. However, an HMAC mechanism must resist to these attacks when part of nodes is compromised. In this section, we present some proposed HMAC schemes for NC-enabled WSNs.

Agrawal and Boneh [5] proposed a HMAC scheme to provide the integrity of the data in network coding. The objective of scheme is to prevent pollution attacks in the case of network coding by the generation of different tags associated to coded packets

In [6], the authors proposed an algorithm which allows the sink node to accept the data with a high probability if the result of the checking of integrity is within an acceptable limit, or to reject the result s 'he's out of bounds. By building a random sampling mechanism and interactive verification, this algorithm offers several methods to securely calculate the HMAC. However, this scheme concerns the security of transmitted data and not the security of generated HMAC tags.

In [7], the authors have proposed a scheme which aims to solve the problem of data authentication by adding several HMAC tags without signature to the payload of the packet. This scheme affects the communication overhead provided by the generated tags.

In [8], the authors presented an efficient aggregation of encrypted data. This solution is designed to provide efficiency and confidentiality in WSNs. The authors propose an additive homomorphic MAC while providing security of aggregated data. The basic idea is to replace the XOR operation by a simple modular addition. This solution is robust against repetitive attacks. Indeed, the security against data pollution attack is not provided since a length key associated with each message.

In [9], the authors proposed a dual-homomorphic MAC scheme based the generation of two different MACs to ensure the integrity of coded packets. The second MAC checks the integrity of the first generated MAC. However, the signature used for the generation of MAC and DMAC for providing security against tag pollution attack is a time-consuming process.

We summarize the related works of secure HMAC schemes for network coding in Table 1, in terms of vulnerability against pollution attacks.

Table 1. Secure network coding schemes

Scheme	Vulnerable to data pollution attack	Vulnerable to tag pollution attack
[5]	X	X
[6]	X	X
[7]	X	X
[8]	X	
[9]	X	X

3. Preliminaries

In this section, we specify some notions on the network architecture used by our proposed scheme as well as the type of attack considered. Finally, we present the homomorphic MAC method.

3.1. Basic Notation

We present in Table 2 the different parameters which will be used in this paper and to formulate our proposed scheme.

Table 2. Notations

Parameter	Description
r	The number of symbols of each packet
s	A finite field size
q,x	The numbers of tags one wants to generate for each vector
l	Prime number
Vect	$\text{Vector} \in F_l^r$
c	Random coefficient
id	Identifier of vector Vect
b	Indicate that the vector Vect is the bth basis vector of the vector space identified by id.
PK	Public key
SK	Secret key
KE	Space key
Mt	q × x matrix over Fq
tg_{ij}	Tag generated for the packet u_i for the vector $Vect_j$
TAG	A tag for the combined vectors Vect

3.2. Network model

For our proposed scheme, we use the model of linear networking presented in Figure 1. A message m is divided into a sequence of vectors: $Vect_1, \dots, Vect_n$ in F_l^r. These vectors are then sent into the network in the form of coded packets u, where there are u vectors for the message m. that is, each packet u represents a vector. The packet u_i is prefixing $\underline{u_i}$ with the ith unit vector represented as:

$$u_i = \left(0, \dots, 0, 1, 0, \dots, 0, \underline{u_{ij}}, \dots, \underline{u_{im}}\right) \in F_l^{r+s} \quad (1)$$

Intermediate nodes combine the packets into a single code as a vector which it subsequently sends across the network. The sink node receives the result and builds the original message after checking its integrity.

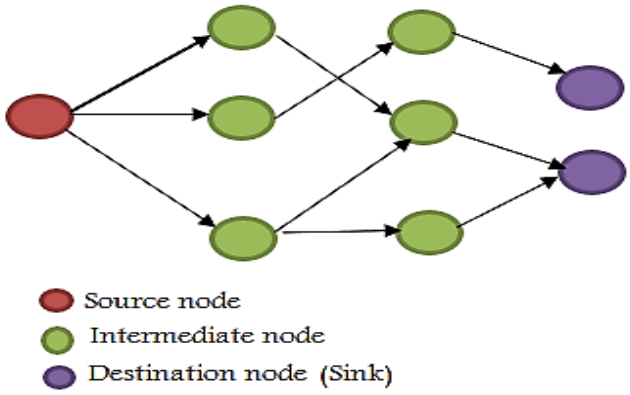

Figure 1. Network model of NC-enabled WSNs where our proposed scheme is applied

3.3. Adversary Model

Pollution attacks [10] aims to add malicious packets in the network which can be used later to start other types of attacks. There are two types of pollution attacks:

- Data Pollution Attack: The mission of an adversary is to insert fake data and to realize the verification of other innocent sensors which causes.
- Tag Pollution Attack: The objective of adversary is to get correct data packets be beaked as false and be isolated by intermediate sinks or nodes, which discard the correct packets.

3.4. Homomorphic MAC

The homomorphic MAC [5] consists of the data of three probabilistic algorithms and in polynomial time (Sign, Combine, Verify):

- $t_u = Sign(k, rid, m_u, id_u)$: the node with the identity id_u ,contributor of the message m_u concerning the relation *rid*, compute a tag t_u for m_u using k as key.
-$t = combine\ ((m_1, t_1, w_1), \ldots, (m_n, t_n, w_n))$:a combiner implements the homomorphic property for the pair (message, tag) in the absence of the key k, which generates the tag t for a combined message $m = \sum_{i=1}^{n} w_i m_i$.
- $Verify\ (k, rid, m, t)$: verify the integrity of the message m in term of rid using the key k and the tag t.

4. Review of Esfahani et al.'s Scheme

Esfahani et al.'s [9] proposed a HMAC scheme by using two types of tags: MAC, DMAC, to ensure the vulnerability against pollution attacks. Each MAC ensures the integrity of the coded packets while each DMAC ensures the integrity of these generated MACs. This scheme comprises four phases: Key setup, MAC generation, Combine, and verify.

4.1. Key Setup

Key Setup Generate in the output a public parameter PK, and sample secret key SK from the key space KE.

4.2. MAC Generation

Before generated the MAC, the scheme generates as input the parameters *(PK, SK, id,Vect, b)*. In the output, the algorithm generates a tag tg for the vector Vect.

4.3. Combine

The input of this step is the public parameter PK, a sequence of triples ($Vect_1$, tg_1, c_1),…, ($Vect_l$, tg_l, c_l), where, for $i \in [1, l]$, tg_i is the corresponding tag of $Vect_i$ under the secret key SK. In the output, the algorithm generates a tag for the combined vector:

$$V = \sum_{i=1}^{m} c_i . Vect_i \qquad (2)$$

4.4. Verify

This step verify the integrity of generated vector V and the tag. The Output is 1 or 0 (reject).

5. Security Flaws of Esfahani et al.'s Scheme

We discuss the security flaws of Esfahani et al.'s scheme. Furthermore, we discover that Esfahani et al.'s scheme cannot ensure vulnerability against pollution attacks.
In the setup step, an adversary A obtains the public and secret keys *(PK, SK)*, and form the parameters (id_i , $Vect_i$). After, A computes, for j=1,...,l:

$$tg_j = MAC(PK, SK, id_i , Vect_j, j) \qquad (3)$$

Finally, the adversary A outputs a tuple (id^*, tg^*, V^*), and A wins the attack if:

$$Verify\ (PK, SK, id^*, tg^*, V^*) = 1 \qquad (4)$$

In conclusion, Esfahani et al.'s scheme [9] is not secure against tag pollution attack.

6. Proposed Scheme

We present an Efficient Homomorphic MAC scheme based on multi-linear space for wireless sensor networks, called E-HMAC, to improve the security flaws of Esfahani et al.'s scheme [9]. Our scheme proposes a new algorithm that supports multi-vector transmission, which supports multi-linear space. The proposed E-HMAC comprises the same process as that Esfahani et al.'s scheme. The details of the four processes are shown below.

6.1. Key Setup

Choose four integer numbers r; s; q; x and a prime number l, where, q,x are the numbers of tags generate for each vector.

Let $G: KE_G \rightarrow M(r+s, x)$ be a pseudo random generator,and $F: KE_F \times x \times (s+q) \rightarrow F_q$ be a pseudo random field. Choose a random SK :

$$SK = (KE_1 , KE_2) \leftarrow KE_G \times KE_F \qquad (5)$$

Output : the public parameters is PK = (l; r; s; q; G; F) and the secret key is SK.

6.2. MAC Generation

In this step, the MAC (PK; SK; id; Vect; b) is generated, where b indicating that the vector Vect is the bth basis vector of the vector space identified by id.
Firstly, compute:

$$SK' = G(KE_1) \qquad (6)$$

,where SK' is the corresponding transition matrix. Let $F_{SK'}$ be the linear mapping : $F_l^{r+s} \rightarrow F_l^x$.

Secondly, compute the dimensional vector a(Vect) :

$$a(Vect) = \sum_{b=1}^{s}(Vect_{r+i} \,.\, F(SK', id, 1, b)), \ldots , \ldots, \sum_{i=1}^{s}(Vect_{r+i} \,.\, F(SK', id, x, b)) \quad (7)$$

Denote y(Vect) the x-dimensional vector, where :

$$y(Vect) = f_{SK'}(Vect) + a(Vect) \quad (8)$$

Let Mt be the following $q \times x$ matrix over F_l :

Mt=

$$\begin{pmatrix} F(SK', id, 1, s+1) & \ldots & F(SK', id, x, s+1) \\ \vdots & \ddots & \vdots \\ F(SK', id, 1, s+q) & \cdots & F(SK', id, x, s+q) \end{pmatrix} = \begin{pmatrix} u_{11} & \cdots & u_{x1} \\ \vdots & \ddots & \vdots \\ u_{1q} & \cdots & u_{xq} \end{pmatrix} \quad (9)$$

Thirdly, compute for i = 1,…,q and j = 1,…,x :

$$tg_{ij}(Vect) = \frac{y_j(Vect)}{u_{ij}} \quad (10)$$

Finally, output the tag :

$$TAG(Vect) = (tg_{ij})_{q \times x} \quad (11)$$

6.3. Combine

In this step, compute and output:

$$TAG' = \sum_{h=1}^{w} c_h . TAG(Vect_h) \quad (12)$$

, where c_h is a random coefficient and w is the number of relay nodes.

6.4. Verify

In this step, we verify the sequence (PK,SK,id,V,TAG). Firsty, compute :

$$SK'' = G(KE_1) \quad (13)$$

, and compute the dimensional vector a(Vect):

$$a(Vect) = \sum_{b=1}^{s}(Vect_{r+i} \,.\, F(SK'', id, 1, b)), \ldots , \ldots, \sum_{i=1}^{s}(Vect_{r+i} \,.\, F(SK'', id, x, b)) \quad (14)$$

Denote y(Vect) the x-dimensional vector, where :

$$y(Vect) = f_{SK''}(Vect) + a(Vect) \quad (15)$$

Let Mt' be the following $q \times x$ matrix over F_l :

Mt'=

$$\begin{pmatrix} F(SK'', id, 1, s+1) & \ldots & F(SK'', id, x, s+1) \\ \vdots & \ddots & \vdots \\ F(SK'', id, 1, s+q) & \cdots & F(SK'', id, x, s+q) \end{pmatrix} = \begin{pmatrix} u'_{11} & \cdots & u'_{x1} \\ \vdots & \ddots & \vdots \\ u'_{1q} & \cdots & u'_{xq} \end{pmatrix} \quad (16)$$

After, compute all the items I:

$$I_{ij}(Vect) = y_j(Vect).u'_{ij} \quad (17)$$

Finally, Check if: $I_{ij}(Vect) = tg_{ij}(Vect)$ for i = 1,…,q and j = 1,…,x.

If all of them hold, output 1 ; otherwise output 0.

7. Security Analysis

The security analysis of our E-HMAC scheme is based on security against the two types of pollution attacks presented in our adversary model in Section 3.

Theorem 1: Without making any changes in the MACs, and suppose that there are m MACs and N neighbor sensor nodes for the adversary A, the probability that the polluted data can succeeds the verification of the neighbor sensor nodes depends on number of keys they have and is not greater than $1/m^N$.

Proof: According to Theorem 1, the probability that the next hops will treat the polluted data as a legal one depends on the number keys that they have. Particularly, if the next hops only have the same key with the compromised node, it will treat the polluted data as a legitimate one. However, for N nodes and m MACs, it happens with the probability of $1/m^N$.

Theorem 2: An adversary A which is one of the N legitimate nodes ants to provide changes in the tg. This attack can be detected by the next hop by the probability of $\frac{1}{2(c+1)}$, where c is a random coefficient.

Proof: The probability of recovering the same shared key between two nodes is $\frac{1}{2(c+1)}$, so the probability of the SK_i from a user i located in next hop i + 1 is $\frac{1}{2(c+1)}$, and this probability would be $(\frac{1}{2(c+1)})^2$, if this key is found in two hops later. However, if the total number of hops between an intermediate node and a base station is (N-i), then the probability that a downstream user receives the same key as SK_i of node i is:

$$Pr = [1 - \frac{1}{2(c+1)}]^{N-i} \quad (18)$$

In fact, the polluted TAG can traverse some hops, before it is detected, by this probability Pr.
Now, A uses id^* in one of the MAC queries, i.e. there exists some b_0 satisfying $id^* = id_{b_0}$. Therefore,

Mt*=

$$\begin{pmatrix} F(SK', id^*, 1, s+1) & \dots & F(SK', id^*, x, s+1) \\ \vdots & \ddots & \vdots \\ F(SK', id^*, 1, s+q) & \cdots & F(SK', id^*, x, s+q) \end{pmatrix}$$

=

$$\begin{pmatrix} F(SK', id_{b_0}, 1, s+1) & \dots & F(SK', id_{b_0}, x, s+1) \\ \vdots & \ddots & \vdots \\ F(SK', id_{b_0}, 1, s+q) & \cdots & F(SK', id_{b_0}, x, s+q) \end{pmatrix}$$

$$=Mt_{b_0} \quad (19)$$

Let $Vect_1, \dots, Vect_s$ be a properly augmented basis of $Vect_{b_0}$ and $TAG_1, \dots, TAG_s$ be their corresponding tags. Let $V^* = (V_1^*, \dots, V_{r+s}^*)$, define :

$$V_0 = (V_1^0, \dots, V_{r+s}^0) = \sum_{i=1}^{s} V_{r+i}^* . Vect_i \quad (20)$$

$$TAG^0 = \sum_{i=1}^{s} V_{r+i}^* . TAG_i \quad (21)$$

Obviously, TAG^0 is a valid tag for V^0 under to the Combine step. In addition, we have, for $1 \leq i \leq s$,

$$V_{r+i}^0 = V_{r+i}^* \quad (22)$$

Since the basis $Vect_1, \dots, Vect_s$ is properly augmented. In conclusion of this analysis, compared to the original HMAC scheme in [9], our proposed scheme E-HMAC protects 50% of the tags from pollution.

Table 3 presents a comparison between our schemes and some schemes presented in the related works, in terms of in terms of vulnerability against pollution attacks.

Table 3. Comparison between our scheme and others in terms of in terms of vulnerability against pollution attacks.

Scheme	Vulnerable to data pollution attack	Vulnerable to tag pollution attack
[5]	X	X
[8]	X	
[9]	X	X
Proposed scheme		

8. Performance Analysis

In this section, we evaluate E-HMAC and compared it with Esfahani et al.'s scheme [9] in term of key storage overhead and communication overhead. Performance results have been performed using WSim and WSNet open-source simulation tools [11].

8.1. Communication Overhead

In our scheme, each source message has M codewords. Let L the number of MAC appended to the source message. Our scheme has the following communication overhead:

$$Communication_{overhead} = \frac{L}{M} \quad (23)$$

Figure 2 shows the communication overhead for E-HMAC and Esfahani et al.'s scheme. Compared to Esfahani et al.'s scheme, E-HMAC achieves 25% less communication overhead. Our proposed multi-linear Space model allows HMAC to achieve this considerable gain in terms of communication overhead.

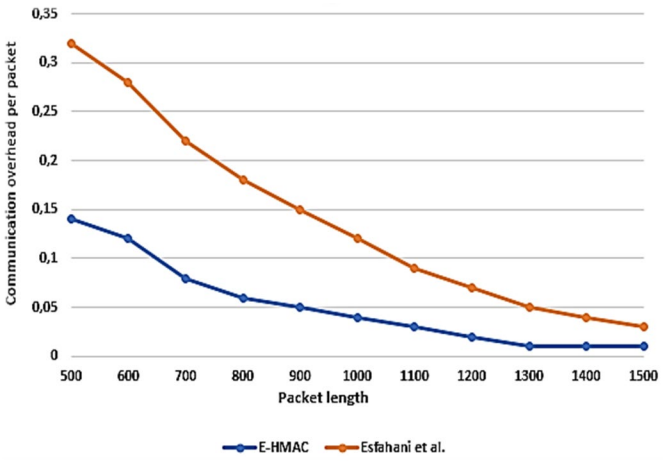

Figure 2. Communication overhead per coded packet

8.2. Key Storage Overhead

We evaluate the key storage overhead of E-HMAC and compare it to the overheads provided by Esfahani et al.'s scheme. Compared to Esfahani et al.'s scheme, E-HMAC ensure almost 50% less key storage overhead at the source node, in terms of the total key storage size required at each source node, which shown in Figure 3.

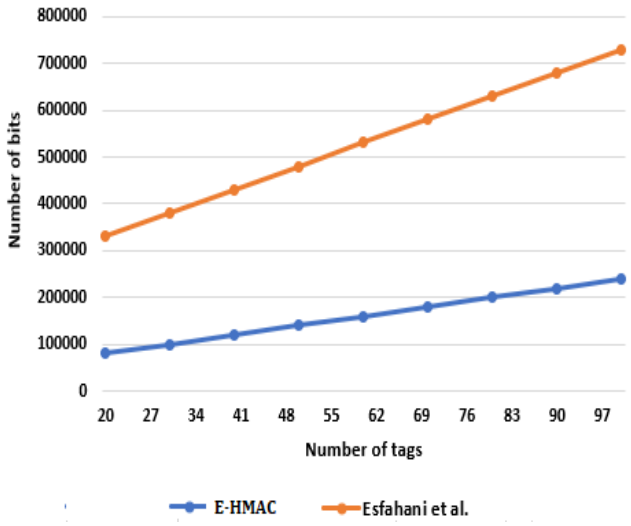

Figure 3. Total key storage size vs number of tags at the source for E-HMAC and Esfahani et al.

9. Conclusions

In Network Coding enabled WSNs, data is split into multiple packets that will be combined and encoded together so that they are transmitted and routed to sink node. The security of these packets presents a challenge in WSNs in order to ensure their integrity. In this paper, we have proposed a new HMAC scheme to ensure an efficient integrity of coded packets in the network. Our scheme is based on multi-linear space, by directly employing mapping over finite fields. The security analysis shows that our scheme is secure against data pollution attack and tag pollution attack. The performance evaluation demonstrates the ability of our proposed scheme to secure the coded packets with a low key storage overhead and communication overhead, compared to Esfahani et al.'s scheme. Extensions to our scheme can be considered to offer additional services such as security of location of sensor nodes in the network.

References

[1] Akyildiz I.F., Su W., Sankarasubramaniam Y., and Cayirci E.''A survey on Sensor networks", *IEEE Commun. Mag.* vol. 40, pp. 102–114, 2002.

[2] Dong J., and Rotaru C.N.,"On the practicality of cryptographic defences against pollution attacks in wireless network coding", *ACM Computing Surveys*, vol. 45, no. 3, 2013.

[3] Walters J.P., Liang Z., Shi W., and Chaudhary V.,"Wireless sensor network security: a survey", *Secur Distrib Grid Mob Pervasive Comput*, vol. 1, 2007.

[4] Boneh D., and Freeman D.M.,"Linearly Homomorphic Signatures over Binary Fields and New Tools for Lattice-Based Signatures", *In: Lecture Notes in Computer Science*, vol. 6571, 2011.

[5] Agrawal S., and Boneh D., "Homomorphic MACs: MAC-based integrity for network coding", *in Applied Cryptography and Network Security: 7th International Conference, ACNS*, pp. 292–305, 2009.

[6] Esfahani A., Mantas G., Monteiro V., and Rodriguez J.,''Analysis of a Homomorphic MAC-based Scheme against Tag Pollution in RLNC-Enabled Wireless Networks", *In: 20th IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks*, 2015.

[7] Lawrence T., Ali I., Christopher T., Li F.,"A bandwidth efficient HMAC-based authentication scheme for network coding", *Journal of Information Security and Applications*, vol. 55, 2020.

[8] Kamal H., and Elbayoumy A., "Implementation of A Homomorphic MAC Scheme in a Transparent Hardware Appliance for Network Coding", *In: 14th International Conference on Computer Engineering and Systems (ICCES)*, pp. 288-292, 2019.

[9] Esfahani A, Yang D, Mantas G, Nascimento A, and Rodriguez J.,"Dual-homomorphic message authentication code scheme for network coding-enabled wireless sensor networks", *Int J Distrib Sens Netw*, vol. 11, pp. 510-251, 2015.

[10] Ning Z., Shi W., Xiao L., Liang W., Weng T.,"A novel approach for anti-pollution attacks in network coding", *Connection Science*, 2021.

[11] Fraboulet A., Chelius G., and Fleury E., "Worldsens: development and prototyping tools for application specific wireless sensors networks", *In 6th ACM/IEEE international conference on Information Processing in Sensor Networks (IPSN)*, pp. 176–185, 2007.