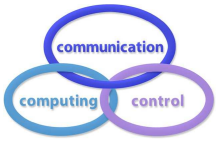
INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 16, Issue: 2, Month: April, Year: 2021
Article Number: 3935, https://doi.org/10.15837/ijccc.2021.2.3935



CCC Publications

SDA-SM: An Efficient Secure Data Aggregation Scheme using Separate MAC across Wireless Sensor Networks

M. Elshrkawey, H. Al-Mahdi

Mohamed Elshrkawey
Department of Information System
Suez Canal University, Egypt
Ismailia 41522
E-Mail: melshrkawey@ci.suez.edu.eg

Hassan Al-Mahdi*
Department of Computer Science
Suez Canal University, Egypt
Ismailia 41522
*Corresponding author: drhassanwesf@ci.suez.edu.eg

Abstract

Securing the aggregated data of the wireless sensor networks (WSNs) is a vital issue to minimize energy consumption and face potential attacks. This paper presents a novel end to end encryption scheme defined as Aggregating Secure Data -Separate MAC (SDA-SM). The importance of the SDA-SM is twofold. First, it separates the secured aggregated data and the message authentication codes (MAC) into two different packets. Second, it transmits these packets in a random separate time-slot according to the scheduling of the TDMA. Moreover, the TDMA applied in the LEACH protocol is modified to adequate to the proposed SDA-SM scheme. The SDA-SM uses MACs to verify the integrity of the aggregated data and uses a sensor protected identifier to authenticate the source of data. The simulation results of the experiments assure the SDA-SM objectives can be achieved with less computation of the communication overheads than earlier techniques. Besides, SDA-SM will be able to accomplish the integrity and confidentiality of accurate aggregated data while saving the energy to prolong the network lifetime.

Keywords: WSN, LEACH, secure data aggregation, homomorphic encryption, Message Authentication Code.

1 Introduction

Wireless sensor networks (WSNs) are employed in various vital applications. These applications may comprise checking the human heart rates, military applications, traffic monitoring, etc. [10]. Actually, the Sensor nodes depend on their limited resources and their predetermined power to survive [12]. However, the functions of the WSN are sensing, aggregation and transmit secured data to the main base station [23]. These functions are faced by the problem of excess energy consumption due to

the transmission of repeated data using the long-range of communication among the sensor nodes [21]. On another hand, the wireless transmissions of the detected data may be prone to eavesdropping, the interception or harming the original data by injecting false data to infect it [15]. So, many scenarios were presented to offer integrity and confidentiality of data as a critical issue. Accordingly, the privacy of data among sensor nodes should be preserved even from the trusted and cooperating sensor nodes [24]. In fact, these problems are managed in the framework of three paradigms: 1) WSN topology control and its routing method for transmitting the sensed data, 2) data aggregation and 3) securing aggregated data before transmission [22].

The WSN is typically self-organized network designed to transmit data in dynamic multi-hops, where vast amounts of power are expended by the sensor nodes during the data transmissions [26]. Clustering topology is managed based on the LEACH protocol to reduce the consumed energy by the sensor nodes [18].

In the data aggregation paradigm, the aggregated sensors accept the various data from the diverse sensor nodes and remove the received data redundancy [8]. Moreover, the aggregated sensors determine the valuable data and transmit it to either the upper aggregator or to the base station. Accordingly, the main objectives of the aggregation process are as follows: First, increasing the WSN energy efficiency by reducing the transmitted messages among the wireless sensors and the base stations [1]. Second, decreasing the exploited cost of deploying and preserving the network. In fact, a variety of data aggregation techniques were presented to shrink the amount of transmitting data over the WSN operations [9]. However, the data aggregation techniques may be suffered from a decline in the QoS metrics, such as latency and accuracy of data.

Finally, the secured data aggregation paradigm aims to decrease the exposure of WSN to malicious intrusions and attacks via introducing the encryption process for securing data. Secure data aggregation can be categorized into two groups, the hop-by-hop encryption and the end-to-end encryption [16]. In the hop-by-hop encryption, the aggregators should decrypt all the received data, combined them using a conforming function and then encrypt the aggregation result before re-sending them to the next-hop. In the end-to-end encryption schemes, each aggregator will receive the ciphertexts from its leaf nodes. The received encrypted data are assembled without decryption and forward it to the next aggregator or to the base station [4]. However, the guarantee of securing the aggregated data is still suffering from destroying or from inserting false data by the adversaries and hackers. This problem can be solved by employing the Message Authentication Code (MAC) to detect false data and achieve data integrity.

In this paper, we present a novel end to end encryption scheme defined as Aggregating Secure Data -Separate MAC (SDA-SM). SDA-SM offers confidential and integrity preserving for the aggregated data in the WSN. The main idea of the proposed SDA-SM scheme is to combine the received data at the aggregators in two separate packets. The first packet contains all secured data and the second packet contains all MACs data. In addition, the MACs are aggregated based on the homomorphic encryption that will be explained later. Two computed and related values defined as the data stamp are inrtoduced. The MAC stamp is created to identify each aggregated MAC to its aggregated and encrypted data packet. Finally, the encryption keys and the secretly shared information are deduced from the Elliptic Curve cryptography. In closing, the main contributions of this paper are as follows:

1. Overcome the problem of the duplicated data due to the overlap of the sensor nodes by improving the management operations of the LEACH protocol.

2. Introduce end-to-end data confidentiality via the implementation of the homomorphic encryption.

3. Increase the aggregated data integrity by forming the homomorphic MAC using data encryption.

4. Overcome the problem of eavesdropping by separating the computed MAC and encrypted data.

The remainder of this paper is organized as follows. Section 2, includes related work. In section 3, the assumptions of the proposed scheme are presented. Section 4 offers the detail of the proposed SDA-SM scheme. The security analyses of the suggested SDA-SM scheme are clarified in section 5.

The simulation results are represented in section 6. Finally, the last section includes the conclusion of this work.

2 Related work

Recently, different schemes have been used by researchers to transmit secure aggregated data across the WSNs. This section reviews the most important previous researches in this area. In [17], EIRDA is implemented on a motionless clustering of a consistent spreading of sensor nodes in each cluster. The operation of EIRDA is initiated when the sink sends the interest packet to the all sensors in the distribution area. The cluster head elect the confidence sensors from the all interested sensors and aggregate the data received from these selected sensors within the cluster and sends them to the sink. Actually, EIRDA objects to offer an effective aggregated data to the sink. in addition, it aims to provide an efficient energy saving scheme and accomplish higher reliability.

In [25], the iPDA accomplishes the integrity of data by employing the redundancy to create two separate aggregation paths to accumulate the disseminated data. The scheme is implemented by allowing each sensor node to divide its sensed data into two different pieces randomly. Each set of pieces are encrypted before transmitting to one of the designated sensor nodes of the aggregation path. In the following, the two aggregated data from two different nodes are received and compared in the base station. The results from two paths are accepted only if the difference between the two sets of assembled data does not contrast with a predetermined value. The iPDA based on excessive operations which increases the number of packets lost and more consumed power.

In [6], the iCPDA scheme was applied to a cluster topology. It is executed on three sequences of actions. Each sensor transmits a primary value to the remaining members in the cluster. This value is used by a node to hide its sensory data. Next, the secret data is transmitted to each member in the cluster. Finally, each sensor node associates its secret data to the incoming secured data and transmits the computed aggregation to the head of the cluster. This protocol suffers from severe overheads which increase periodically when the cluster size is increased. In [27] and [13], authors propose schemes based on the elliptic curve discrete logarithm. However, in [27], the presented scheme isn't able to meet the security requirements such as confidentiality and source authentication. In [13], the authors present a hierarchical data aggregation scheme to ensure confidentiality. Each cluster is assigned to its public key. The encryption is performed when the messages data from all cluster members have arrived. Hence, the transmission of unprotected data may affect its integrity.

In [7], the authors implemented a homomorphic hash for the encryption of the combination of data. They replace the XOR process with a modular addition. However, this proposal is restricted to work only against passive attacks. Finally, in the [14], the authors propose an Aggregated MAC (AMAC) that implemented based on MAC to offer authentication. The implementation based on AMAC aimed to diminish the cost of the transmission deserved by MAC. The authors offer a cluster-based scenario by applying MAC to decrease the number of bits used for authentication during the communication.

In this paper, the EIRDA, iPDA and MAC are elected as the most popular schemes that were investigated and compared by the researchers for testing the energy consumptions and measuring the security aspects for their proposed schemes in the WSN. However, the proposed scheme is distinguished than these schemes in a two main advantages. firstly, it separated the aggregated data packet than the aggregated MAC packets and each of them transmitted in different path. second, it consumes less energy when compared by the mentioned schemes as clear in the simulation section.

3 Cluster WSN description

As shown in Figure 1, the WSN includes a base station (BS) that gathering information to the end applications and operates as a sink node. The coverage area of the WSN is divide into clusters N_C. Each cluster includes N_S sensor nodes. Each sensor is referred to as Cluster Member (CM) with a unique identifier. Let ID_i^j be the identifier of the CM i in the cluster j, where $i = 1, 2, 3, \ldots, N_S$ and $j = 1, 2, 3, \ldots, N_C$. In each cluster, one of its CMs is selected to be cluster aggregator (CA). The CM sensors periodically collect data from their monitored area into their data buffer, then transmit

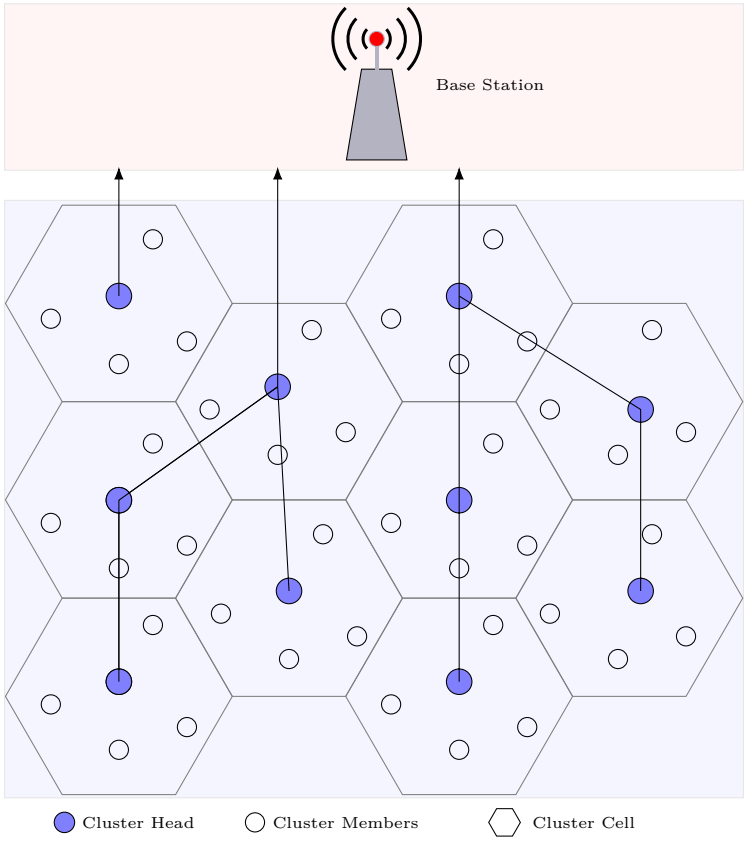


Figure 1: Network model of the SDA-SM

the data to their CA. The CA sends the collected data to the BS. The WSNs are managed based on the modified LEACH protocol which has high significance in clustering routing protocols [5]. The operations of this protocol are implemented in successive rounds. Each round includes a setup phase and a steady-state phase. In the setup process, the clusters are formed and one of its sensors is employed as CA. Before the installation process of the WSN, initial keys are created in the offline state at both the CMs, CA and BS. These keys are used later to generate secret shared information. The secret shared information will be deployed to create the required session encryption keys.

4 The proposed SDA-SM scheme

In the proposed SDA-SM scheme, the initial keys consist of a private key K_R and a public key K_U . The value of K_R is manually selected in a random manner and stored at both sensors and BS. The value of K_U will be generated according to the Elliptic Curve Cryptography (ECC) [3]. The cryptosystem of the ECC is based on the six parameters of the elliptic curve domain over H_p . Where H_p is a finite field consists of the set of integers modulo a large prime number p.

To generate K_U , three out of these six parameters are used to produce it. These three parameters are denoted as p, $G(g_x, g_y)$, and n , where $G(g_x, g_y)$ is a base point on the elliptic curve $E(H_p)$ and n denoting the order of base point $G(g_x, g_y)$ (i.e., the number of different points over $E(H_p)$. The mentioned three parameters are combined with the value of K_R to create K_U using algorithm 1 as follows.

At this end, each sensor node includes the tuple $\left(K_R, ID_i^j\right)$. However, the tuple (K_R, K_U) is stored in the BS. After generating the initial keys, the proposed SDA-SM secures the aggregated data in four successive processes as depicted in Figure 2. These processes are setup process, encrypt-sign process, aggregation process, and verification process. In the setup process, the required of the session keys will be created. The encrypting process includes the implementation of the secure technique for the sensed data and generating the encrypted tag for each CM. The aggregation process is used for

Algorithm 1 Generating the initial keys

1: Input a large prime number and base point order p and base point order n.
2: Generate the base point $G(g_x, g_y)$ using ECC.
3: $K_R \leftarrow random(1, n-1)$
4: $K_U \leftarrow K_R.G(g_x, g_y)$
5: Output**:** the vector K_U with two components.

aggregating data of all sensors node to their aggregator within each cluster. In addition, a single MAC based on the homomorphic property is created within each cluster [11].

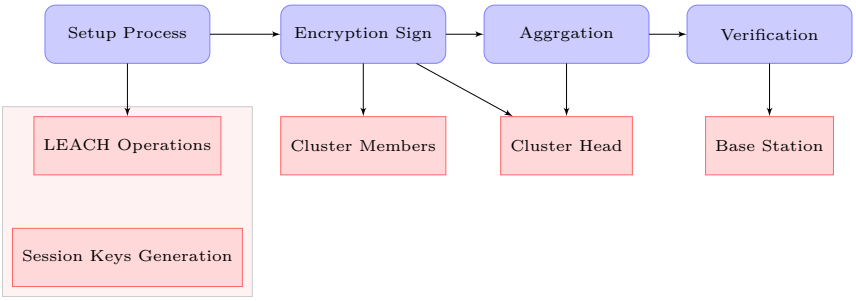


Figure 2: General processes of the SDA-SM scheme

Finally, the arrival of the secure aggregated data to the BS should be verified in the verification process. In the following subsections, more details about the operations of each process are clarified. Also, the details of the proposed contributions in the SDA-SM scheme are explained and illuminated.

4.1 Setup process

This process is implemented at both the BS and sensors. First, the BS generates the two points S_b and A_b based on the value of the key K_U as follows.

$$S_b = r \cdot K_U = (S_x, S_y) \tag{1}$$

and

$$A_b = d \cdot K_U = (A_x, A_y). \tag{2}$$

Where $r, d = random(1,\ n-1)$. The points S_b and A_b are used later to generate the session keys at the BS. Second, the BS generates the shared secret information points R and B using the ECC as follows:

$$R = r \cdot G(g_x, g_y) = (r_x, r_y) \tag{3}$$

and

$$B = d \cdot G(g_x, g_y) = (b_x, b_y). \tag{4}$$

The BS broadcasts the points R and B to the all sensors. Upon the shared secret information points R and B are received by the sensors, they start to build the two points S_S and A_S as follows:

$$S_S = K_R.R = K_R.r.G(g_x, g_y) \tag{5}$$

and

$$A_S = K_R.\ B = K_R.d.G(g_x, g_y). \tag{6}$$

One of the main contributions of the proposed SDA-SM scheme is that the value of pairs (S_b, A_b) and (S_s, A_s) are equal. This can be verified using (1), (2) and the value of K_U as follows:

$$\begin{aligned} S_S &= K_R.r.G(g_x, g_y) = r.(K_R.G(g_x, g_y)) \\ &= r.K_U = (r_x, r_y) = S_b \end{aligned} \tag{7}$$

and

$$\begin{aligned} A_S &= K_R.d.G\left(g_x, g_y\right) = d.\left(K_R.G\left(g_x, g_y\right)\right) \\ &= d.K_U = \left(A_x, A_y\right) = A_b \end{aligned} \tag{8}$$

As a result, the generation process of the session keys at the BS using the points S_b and A_b gives the same results at the sensors based on the point S_b and A_b. Finally, the generation process of the session keys K_D, K_M and K_{ID} is started at the BS and sensors using the points (S_x, S_y) and (A_x, A_y). The K_D, K_M and K_{ID} are used to cipher/decipher data, digital signature and identifier of each sensor respectively. For all sensors and the BS, the values of , and K_{ID} are obtained using the well-known key derivation function $KFD\left(x, y\right)$ [2] as follows:

$$K_D = KFD\left(\left(S_x, A_x\right), K_R\right), \tag{9}$$

$$K_M = KFD\left(\left(S_y, A_y\right), K_R\right) \tag{10}$$

and

$$K_{ID} = KFD\left(\left(K_D, K_M\right), K_R\right) \tag{11}$$

Generally, these keys can be changed from session to session by changing the base point $G\left(g_x, g_y\right)$.

4.2 Encrypt-sign process

In this section, the plain message ($pMsg$) encryption procedure and CM signature tag will be explained as follows. First, in the proposed SDA-SM scheme, the $pMsg$ at each CM is divided into chunks $m_l, l = 1, 2, 3, \cdots, N$ where m_l has length of L bits . The encryption procedure is started by masking each chunk using A_y stream of bits as $M_l = m_l + A_y$. The masked chunk is encrypted through 32 stages of the Unbalanced Feistel network [19]. A Feistel network is a general method of transforming any function into a permutation. It was proposed by Horst Feistel in his design of Lucifer [20]. The Feistel function, F takes M_l and K_{MSG} as inputs and the output is cipher message ($cMsg$). Where the value of K_{MSG} is generated at every CM based on its identifier ID_i^j as follows:

$$K_{MSG} = K_D \otimes PKG_i^j \tag{12}$$

where the symbol $\otimes$ denotes the XOR logical operation and PKG_i^j is a packet-guide used to hide the value of ID_i^j. The value of PKG_i^j is given as:

$$PKG_i^j = ID_i^j \otimes K_{ID} \tag{13}$$

The proposed SD-SMA scheme encrypts the $pMsg$ at each sensor through 32 subsequent stages using algorithm 2.

Second, the sensor tag signature is generated as follows. Each CM i of the cluster j computes its tag signature T_i^j based on the values of K_M , ID_i^j and $cMsg_i^j$ as follows:

$$T_i^j = \left(cMsg_i^j + ID_i^j\right) \otimes K_M \tag{14}$$

Finally, each CM i of cluster j sends the triplet $\left(cMsg_i^j,\ PKG_i^j,\ T_i^j\right)$ to its CA as shown in Figure 3.

4.3 Aggregation process

One of the most important contributions in this scheme is the separation among the aggregation of the encrypted messages and the message authentication code. Upon the CA of cluster j receives the triplet $\left(cMsg_i^j,\ PKG_i^j,\ T_i^j\right)$ from its N_s active CM, it starts to compute the total message authentication code MAC_T^j as follows:

$$MAC_T^j = \sum_{i=1}^{N_S} T_i^j \tag{15}$$

Algorithm 2 Encryption process of *pMsg*

Input $pMsg$
Compute K_{MSG} using 12 and 13.
Divide $pMsg$ into chunks M_l, $l = 1, 2, 3, 4$.
Set $K^0_{MSG} \leftarrow K_{MSG}$, $w \leftarrow 1$
while $w \leq 32$ **do**
    if w is odd **then**
        $K^w_{MSG} \leftarrow K^{w-1}_{MSG}$
    else
        $K^w_{MSG} \leftarrow 2 << K^{w-1}_{MSG}$
    end if
    for $j \leftarrow 1$ to 4 **do**
        $M^{w+1}_j \leftarrow F\left(M^w_j, K^w_{MSG}\right)$
    end for
    $w \leftarrow w + 1$
end while
$cMsg = M^{w+1}_1 + M^{w+1}_2 + M^{w+1}_3 + M^{w+1}_4$

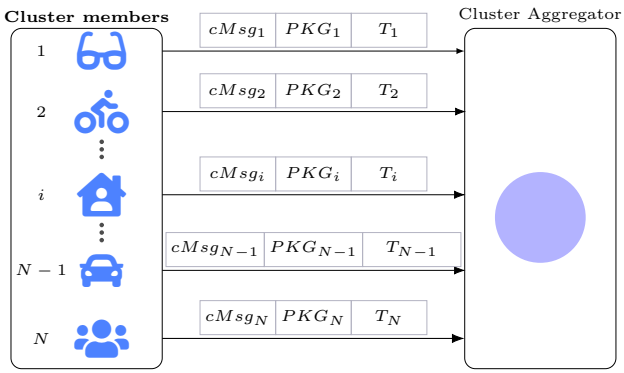


Figure 3: Cluster members send messages to their cluster head

Once the value of MAC_T^j is calculated, the CA of cluster j creates two stamps known as data stamp ST_{DATA}^j and MAC stamp ST_{MAC}^j as follows:

$$ST_{DATA}^j = ID_{CA}^j \otimes K_{ID} \tag{16}$$

and

$$ST_{MAC}^j = ID_{CA}^j \otimes (2 << K_{ID}) \tag{17}$$

Where ID_{CA}^j denoting the CA identifier of cluster j . As shown in Figure 4, the CA of cluster j formulates all the received triplet $\left(cMsg_i^j, PKG_i^j, T_i^j\right)$, the MAC_T^j and the stamps in two separate packets as follows:

1. For each CM i of the cluster j , the pair $\left(cMsg_i^j,\ PKG_i^j\right)$ is filled into the aggregated data packet one by one followed by ST_{DATA}^j at the end.
2. The values of MAC_T^j and ST_{MAC}^j are stored in the aggregated MACs packet.

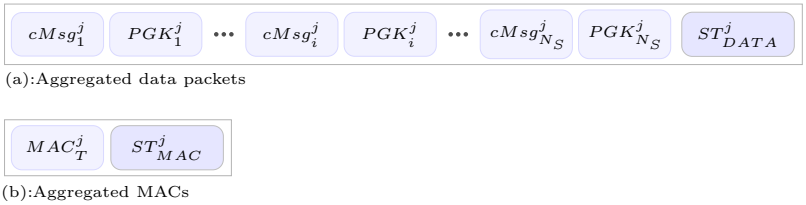


Figure 4: Aggregated packets for CA of cluster j

4.4 Verification process

For each session, the verification process is accomplished at the BS using the cluster aggregator identifier ID_{CA}^j and the session keys, K_d, $K_M$and K_{ID} . These keys are obtained using equations (9)-(11). First, for CA of cluster j , the received aggregated data and MACs packets at the BS are matched with their correspondence cluster aggregator j as follows:

1. The BS computes the data stamp $ST_{DATA-BS}^j$ and the MAC stamp using (16) and (17) respectively.
2. The BS compares the values of $ST_{DATA-BS}^j$ and ST_{MAC-BS}^j with the received values ST_{DATA}^j and ST_{MAC}^j respectively. If they are equal, this means the received aggregated data and MACs packets are verified with cluster aggregator j . Otherwise, the aggregated data packet and its MAC are neglected.

Second, upon the aggregated data and MAC packets are verified with cluster j , the BS starts to decrypt each packet-guide concatenated to each message in the aggregated data packet of cluster j to get the identifier of CM associated with that message and then decrypt the message itself. The process is clarified in algorithm 3.

4.5 Steady-state phase of the LEACH protocol

Each round time in the LEACH protocol consists of a setup phase period and steady-state phase period. The setup phase period includes the WSN initialization and scheduling for the transmission process. In the steady-state phase, time is divided into slots. Each set of slots makes up a frame as depicted in Figure 5. The transmission process is implemented according to the time division multiple access (TDMA) schedule. The scheduling process is customized by the CA of the cluster before notifying all remaining members in the cluster. According to the proposed SDA-SM scheme, both of

Algorithm 3 Verification process

: data packet and MACs packet.
for $j \leftarrow 1$ to N_C **do** //*iterate for all active CA*
  $MAC^j_{T-BS} \leftarrow 0$
  for $i \leftarrow 1$ to N_S **do** // *iterate for all active CMs j*
    $ID^j_i = PKG^j_i \otimes K_{ID}$
    Calculate the tag signature T^j_i using 14.
    $MAC^j_{T-BS} \leftarrow MAC_{T-BS} + T^J_I$.
  end for
  if the received $MAC^j_T = MAC^j_{T-BS} =$ **then**
    Accept the aggregated packet.
    for $i \leftarrow 1$ to N_S **do** // *iterate for all active CMs j*
      Decrypt the $cMsag^j_i$ using ID^j_i and PKG^j_i
    end for
  else
    Reject the aggregated data packet and its MAC.
  end if
end for

the aggregated data and MAC are stored in separate packets. So, the proposed contribution for the steady-state phase is dividing each time slot within a frame into two mini-slots as shown in figure (5). The min-slot period is sufficient to transmit an aggregated data or MAC packet. This division will allow each aggregator to send its two packets in scattered min-slots (i.e., random assignment of the mini-slots). When the separation process is added to the security operations performed in each CM and CA, a robust confidently and integrity will be deduced for the aggregated data.

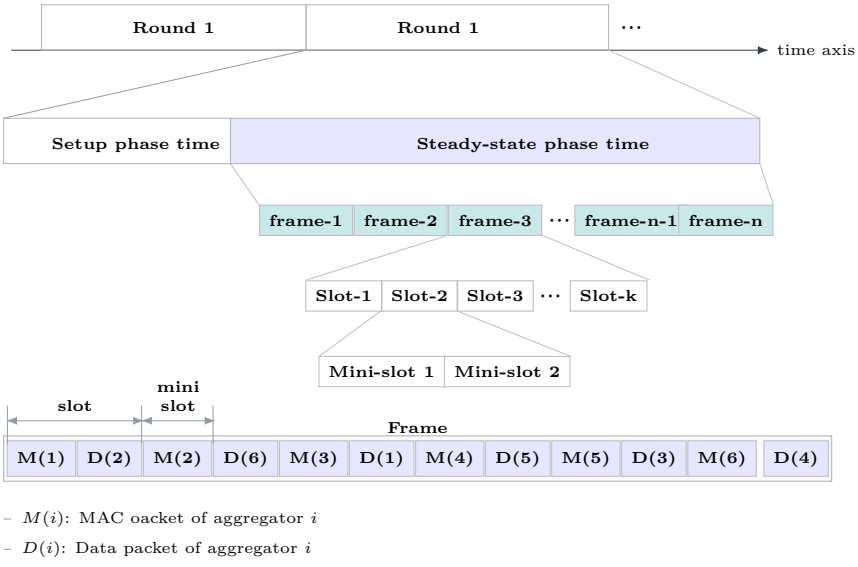


Figure 5: The proposed TDMA for the aggregators transmission

5 Security analyses

In spite of the multi-hops is an implemented approach to minimize energy consumption in WSN, it deduces a multiple of the security challenges. So, the implementation of end-to-end security services becomes obligated and required. In the following subsections, the roles of the SDA-SM scheme in this context will be revealed. The SDA-SM scheme is analyzed according to the security objectives such

as confidentiality, integrity, and authentication. Additionally, the SDA-SM as a robust scheme will be investigated against different types of mentioned attacks.

5.1 Data confidentiality

AAs aforementioned, each CA sends its aggregated data and MAC packets in two separate mini-slots within a frame with length Z slots (i.e. $2Z$ mini-slots). Assume that the attacker succussed to obtain the aggregated MACs packet of the cluster aggregatorwithin a certain frame.

Step 1: He tries to determine the CA who owned the captured aggregated MACs packet by computing the value of ST^j_{MAC}, $(j = 1, 2, 3, \ldots, N_c)$ for each CA and compare the result with stamp MAC within the captured aggregated MACs packet using equation (9). To do this he needs all the possible values of the both aggregator identifier ID^j_{CA} and the private key K_{ID} . The length of ID^j_{CA} is 32 bits (i.e., 4 bytes) while the length of K_{ID} is 34 bits. As a result, the number of trials needed for this step is given as

$$ST_{trials} = \sum_{j=1}^{N_S} j \times 2^{66} \tag{18}$$

The probability of success, $P_{St-Sucess}$, in this step is given as

$$P_{St-Sucess} = \frac{1}{ST_{trials}} \tag{19}$$

Step 2: Assuming that, the attacker succussed in step 1. In such case, he tries to determine the correct aggregated data packet with total tag signatures equal to the MAC^j_T. The tag signature of each message within the aggregated data packet were calculated based on the values of K_M , ID^j_i and $cMsg^j_i$ as illustrated in (14) and (15). All these parameters are unknown to the attacker except the message $cMsg^j_i$. The number of trials needed to compute the tag signature of CM i in cluster j within the aggregated data packet is given as trials. The total number of trials for the whole aggregated data packet is given as

$$T_{agg-MAC} = \sum_{i=1}^{N_S} i 2^{64} \tag{20}$$

The attacker repeats this process to all the aggregated data packets with the transmitted frame. Consequently, the final number of trails is given as

$$T_{MAC} = \sum_{i=1}^{N_S} i \times 2^{64} \times F \tag{21}$$

The probability of success in this step is given as

$$P_{MAC-Success} = \frac{1}{T_{MAC}} \tag{22}$$

Finally, the probability of attacker failure is given using (22) and (22) as follows:

$$P_{failure} = 1 - P_{St-Sucess} \times P_{MAC-Success}$$

Moreover, assume the attacker succeeded in getting some of the messages and some of MACs; he will collide by both of the encryption keys used in securing the stamps. Therefore, the attacker trails may approximately tend to zero. Thus, the confidentiality of the SDA-SM scheme is accomplished regardless of any further actions.

5.2 Data Integrity

In the SDA-SM scheme, integrity is built to accomplish the second step of security. Firstly, it is based on homomorphic MAC. This MAC is computed based on three unknowns for the attacker. They are the encrypted message, the embedded identifier of each sensor and the MAC key K_M. Assuming

an adversary is succeeded in catching any encrypted message, he must know the session keys (I.e. K_D, K_M, K_{ID}) that are used for encrypting the message components. However, these keys K_D, K_M and K_{ID} built based on secret information that is embedded within the sensors and base station before the deploying of the wireless network topology. However, these keys can be changed every round if required by changing the two points of secret information that are ciphered and deciphered based on the values of K_U and K_R which are embedded within the sensors and base stations without exposing to any mutual propagations. Moreover, the attacker must know the MAC matched to each catches message. The existence of the related MAC is obligatorily required to authenticate the catched message. On the other hand, the encrypted sensor message is required as a part of the digital signature computed for each sensor within each cluster. As a result of the two last equations, the probability of getting the required unknowns is tending to zero. Again, the essential separation of the aggregated data packets and aggregated MAC packets will make an adversary tries cannot effectively copy specified messages with their related MACs. Furthermore, if an adversary succeeded in performing any alteration in the message through transmission, the base station will be able to detect this alteration based on the strange identifier presented by an adversary through this process of the sensor authentication. Because each encrypted message is assembled in the transmitted packet followed by the encryption of the identifier of its sensor (PKGi). So, the attacker will fail in performing any effect on the integrity of the message. Thus, the proposed SDA-SM scheme preserves robus integrity.

5.2.1 Unauthorized aggregation

Both of ID_i^j and K_{ID} are unknowns to an adversary. In the SDA-SM scheme, the aggregations of the encrypted data and the encrypted MACs are performed in two separate packets based on data that already encrypted and tied by two different computed values defined as the packet stamps and MAC stamp. Both of aggregated data and MAC are tied by two different computed values defined as the packet stamps and MAC stamp. Hence, these stamps should be decrypted to identify each aggregated packet to its MAC. However, these stamps are the encryptions of the aggregator identifier with two different keys $\left(K_{ID},\ ID_{CA}^j\right)$ that mentioned before. Desides, each sensor data is concatenated with a packet guide that is encrypted based on the embedded identifier of its sensor ID_i^j and the encrypted key K_{ID}. So, when an attacker tries to perform the false aggregation, he must know these different session keys, K_{ID}, ID_{CA}^j, and ID_i^j for each participated sensor and the applied K_{ID}. Since these secret keys and identifiers are disappeared relative to an attacker, he will not able to perform the required and right aggregations. On the other hand, these keys can be changed every round if required. So, if these keys are obtained to him, it will become useless. Even though, any presented forged data to the aggregated data packet will be detected through the MAC verifications that are propagated on a different channel. So, the proposed scheme is more to fight against unauthorized aggregation when compared with other schemes. iPDA as an instance, the data are sliced and transmitted in different paths. But the encrypted data are decrypted within the aggregator nodes during their transmission. So, it is more exposed to false aggregation. In addition, for both of the EIRDA and AMAC, the session keys are exchanged among the base station and the sensor nodes. Moreover, both the encrypted data and encrypted MACs are grouped in a single packet. So, both of them may be exposed to unauthorized aggregation when these session keys or messages are obtained by the attackers.

5.2.2 Malleability

The Malleability permits the alterations of the ciphered data regardless of inevitably know its substances or its source. The alteration may be performed simply by embedding any data similar to ciphered data. So, the SDA-SM scheme offers end-to-end integrity by allowing the sink to confirm the singular data by authenticating every source node through its identifier that is defined in the base station. So, if ciphered data of a specific source is changed, the verification process will be failed and this source of aggregated data will be excluded. Moreover, the sink will jump to the detection procedure about the malicious node. In the SDA-SM, the detection procedure about the malicious nodes is performed in the sink through the identification process. On other hands, any alteration in transmitted ciphered data will be detected through the MAC verification procedures.

Table 1: Simulation parameters

Parameter	**Value**
Number of Sensor Nodes	50-450 nodes
Range of Transmission	30m
Dimensions of Work Area	$450 \times 450\ m^2$
Transmission Power	0.650, 0. 125 mw
Received Power	0.375, 0.75 mw
Preliminary Energy	7.2 J
Noise Floor	-80 dB
Simulation Time	400 sec.

5.2.3 Node compromise

In the SDA-SM, if the malicious node can capture one of the aggregators, it will not be able to access the plain contents of its members. Originally, there are no plain data appear in the aggregators. The only side effect may be performed by the adversary is the editing of the encrypted message of the aggregator. However, the compromised node should also be able to get the encrypted packet of the separate MAC that transmitted on a different channel. This separation presents a different guide than all the presented methods.

5.2.4 Replay Attack

An adversary tries to acquire a portion of legitimate traffic for a period of time before resending again to impose an aggregation error to affect the entire of the aggregated results. However, an adversary tries were based on valid packets already transmitted to the base station However, an adversary will not be able to perform his try the legitimate encrypted and aggregated packet without obtaining its related MAC and different session keys. Because, the MAC transmitted on a different channel is unknown for the attacker. Even though, as very lowest probability to get both of aggregated packet and its own MAC. AS the modified data sent to the base station, the received cipher data will become unknown. Because the applied encryption keys were changed in the next round. So, adversary operations are simply detected.

6 Simulation results and performance analysis

MATLAB 2015a is employed to simulate the SDA-SM scheme and the compared schemes. The 2D Elliptical Gaussian distribution was applied. The WSN topology was embodied by the number of sensors between 50 and 450. The deployed work area was simulated in square region 450×450 m^2. The preliminary energy is $7.2J$ for each CM. Generally, all simulated parameters that are implemented to investigate the SDA-SM are shown in Table 1. To minimize the error of the comparisons, the experiments of the SDA-SM simulation and compared schemes iPDA, EIRDA and AMAC were performed regularly for 20 times. Accordingly, the results of the average value of the experiments were measured for both terms of communication overhead and energy consumption.

6.1 Communication overhead

To verify an evident comparison among the SDA-SM and the compared schemes, the communication overhead evaluated by the summation of the sending bytes from all the sensor nodes during the aggregation phase. The communication overhead was determined by the excess number of bytes that occurred as a result of the aggregation process for the ciphertexts and the MAC. The excess

bytes are the bytes embedded by the sensor nodes to produce both of the ciphertexts and the MAC before transmission. The count of the participated sensors was considered during the measure of the aggregation performance. The results are shown in Figure 6. The figure reveals the amount of excess data that was generated by all schemes to perform the aggregation process for the same real sensing data. It is clear that the amount of excess data that was required by the SDA-SM scheme is the minimum amount of the overhead data when compared by the other schemes for the same real sensing data. Besides the fact of the excess data, when the number of sensor nodes increased, each aggregator should be able to aggregate data from more sensors. So, the waiting time will be increased for the aggregators and the sink until all aggregated data are received. So, snapshots of the sink waiting times were measured at 100 sensors for all schemes. It is noticed that the sink waited approximately for 1.2 sec to receive the aggregation of the encrypted data using iPDA, 0.9 sec to receive the aggregation of the encrypted data using EIRDA, 0.70 sec using AMAC while it waited for only 0.2 sec for the SDA-SM scheme. Actually, the SDA-SM enhancement obtained due to the reduction in the sizes of MAC and the data packets. Furthermore, an enhancement is deduced in bandwidth utilization due to the same reason.

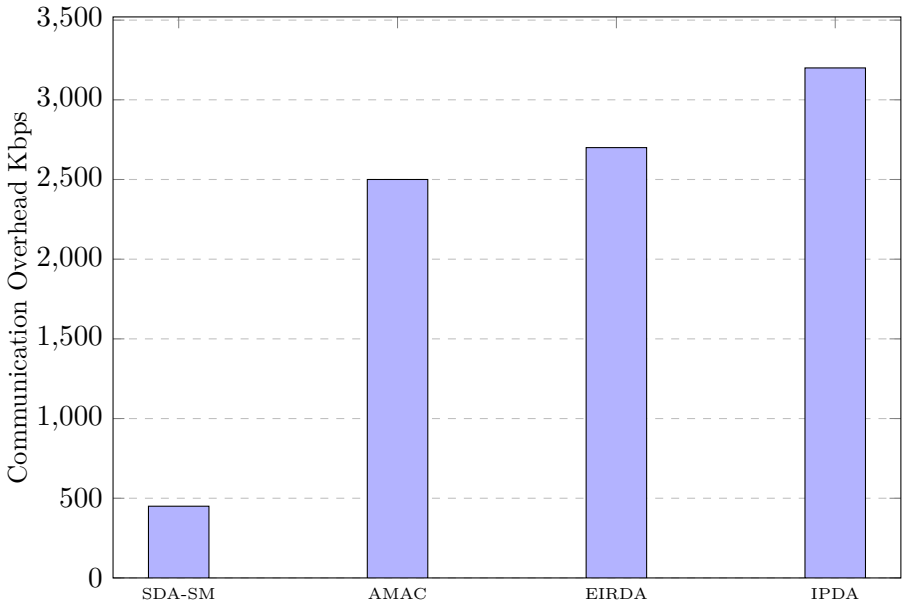


Figure 6: Communication overhead

6.2 Energy consumption

In the following subsections, Energy Consumption by the applied topology and the sensor nodes through communication are introduced. In addition, a comparison among the energy consumed per packet in the proposed scheme and some of the previous schemes is illustrated

6.2.1 Energy consumption by sensor nodes

In the WSNs, energy consumption incarnates the vital issue. Most of the energy consumed by the sensors of the wireless network occurs due to the process of the transmission of the packets. When the amount of the transmitted data is increased, the energy consumption will be increased causing faster end in the life of sensor nodes. So, the measures of the average of the consumed energy by the sensor nodes of the implemented topology are revealed in Figure 8 for the four compared schemes. The simulations are repeated and measured at the different number of sensors 150, 300 and 450 as exposed in Figure 7. Generally, the averages of the consumed energies are increased as the numbers of the sensor nodes are increased. However, the SDA-SM scheme provides a great enhancement in the depletion of energy by comparing it to the other schemes. This enhancement is deduced due to the

reduction in the operation cost before transmitting data as compared with the operation performed in the other schemes for the same amount of the sensed data. Also, the SDA-SM performs its operation using the lowest of the communications overhead among all schemes. Therefore, the average quantity of the energy depleted by the sensor nodes is greatly reduced in the SDA-SM scheme. Consequently, it will provide a momentous enhancement in the network lifetime when compared with the other schemes.

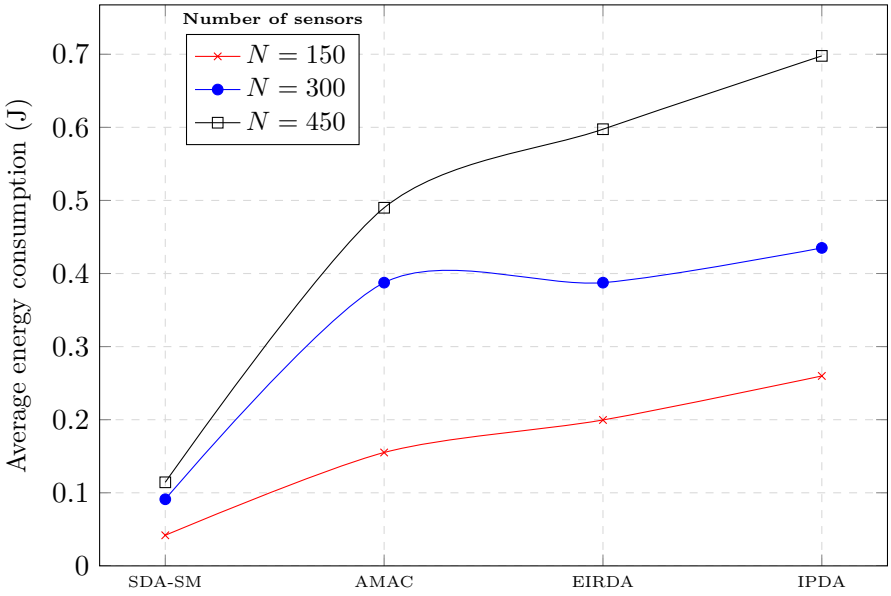


Figure 7: Average of the consumed energy by the sensor nodes

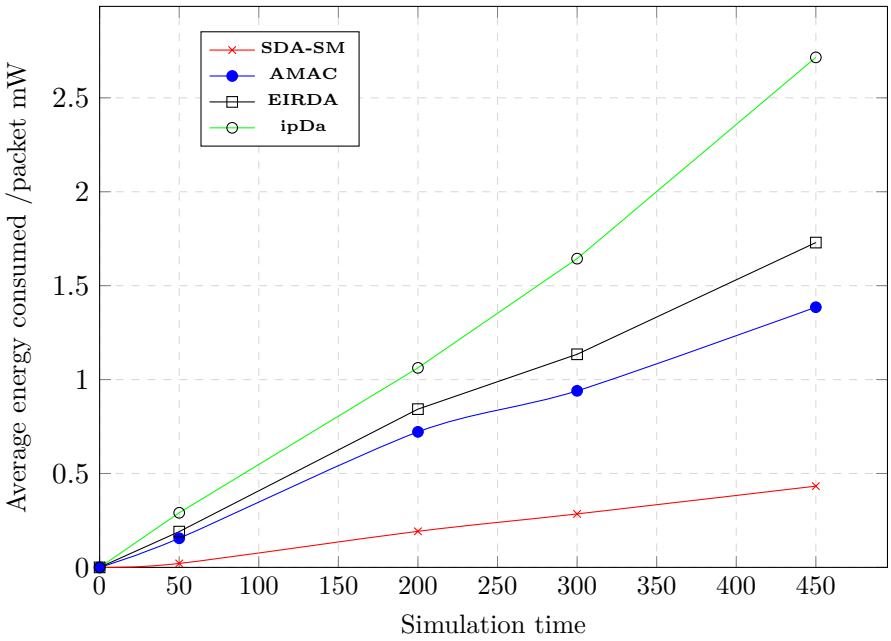


Figure 8: Comparison of the average energy consumed per packet

6.2.2 Average energy consumption per packet at base station

These experiments were implemented to measure the average energy consumed per packet in each round over a different number of nodes. The computations performed according to the simulation

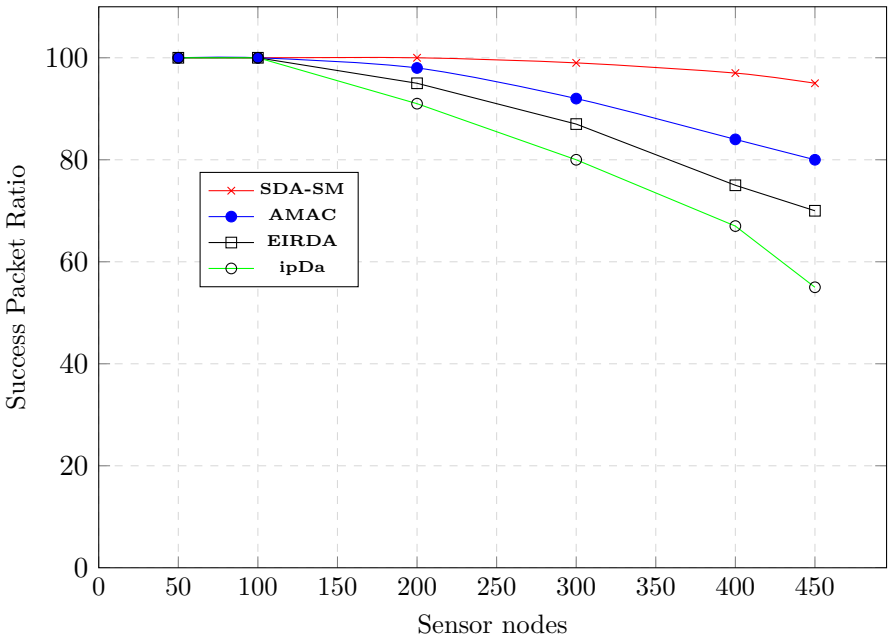


Figure 9: Ratio of success received packet at different number of nodes

parameters table. The average energy consumed per packet was computed in the following two steps. Firstly, the energies consumed by all sensor nodes for transmitting the packets from their source to the sink were summed. Secondly, the result of summation was divided by the number of success received packets. The total energies consumed per packet for the proposed scheme and other compared schemes are revealed in Figure 8.

In general, the simulation results of the compared schemes indicate the increase in energy consumption per the packets as the number of nodes is increased. Truly, the increase in the number of nodes means the increase in the number of transmitted packets and consequently an increase in the depleted power through the transmitting and receiving operations. On the other hand, some of the aggregator nodes may die which leads to more increase in power consumption due to the re-configuration of the topology and change in the aggregator nodes. However, the figure shown the proposed scheme SDA-SM accomplishes the minimum energy consumption per the packets when compared by the other schemes. The enhancement is deduced due to two main reasons. First, the modified aggregation process clarified in subsection 4.3. Second, the TDMA scheduling method applied in subsection 4.5.

6.3 The throughput at the base station

These experiments were implemented to compare the average Ratio of Success Packets Received at the base station for the SDA-SM and the all compared schemes including EIRDA, iPDA, and AMAC. The simulations of all compared schemes were run on the number of nodes 50, 100, 200, 300 and 450. Moreover, each simulation was run at an average of 10 times. The results of these experiments are offered as exposed in Figure 9. The marks on the different curves indicate the percentage of the success packets received at the base station for each method. It is clarified that the percentage of the success packets will be lessened as the number of nodes rises for all schemes. Through the exchange of the packets among the WSNs, some of the cluster heads may exhaust their energies. So, some of the packets may be lost during their transmission to their base station. Therefore, the numbers of the successful packets received may be lessened if the amount of the energy of some sensor nodes especially the aggregators is exhausted.

Accordingly, some of the aggregated packets may be lost. However, the proposed scheme accomplishes the maximum success ratio for the received packets when compared by all other schemes. This enhancement is accomplished due to the minimum energy consumption of the proposed scheme through the minimum cost that is needed for the aggregation processes and security operations. Also,

the reduction in the energy of the proposed scheme is deduced from using the modified LEACH presented in [5]. In this modified LEACH, the enhancement is achieved due to the method used for the Cluster head selection. The selection method confirms a well-adjusted generation of aggregators among all clusters. This constancy leads to a well-adjusted choice causing the regularity in energy dissipation.

7 Conclusion

In this work, a novel scheme SDA-SM is offered to accomplish the needed security for the exchanged aggregated data among the sensor nodes and the base stations. In addition, the SDA-SM scheme aims to extend the lifetime of the WSNs. The SDA-SM is based on three different contexts to achieve the required objectives. Firstly, it develops the TDMA scheduling method used by the LEACH protocol to decrease the amount of the lost energy during the transmission of packets among the sensors. Secondly, it is applied to an additive homomorphic on the encrypted tags to perform the required aggregation for the total MAC. Finally, the separation between the aggregated MAC represented in a separate packet and the aggregated data represented in other packets to get robust security implementation. The SDA-SM scheme is discussed and analysed to illuminate the achievement of the security objective and its fight to different attack methods to clarify its forceful in succeeding its aims.

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Alduais, N. A. M.; Abdullah, J.; Jamil, A.; Audah, L. (2016, October). An efficient data collection and dissemination for IOT based WSN. In *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, IEEE, 1-6, 2016.

[2] Barker, E.; Chen, L.; Davis, R. (2018). Recommendation for Key-Derivation Methods in Key-Establishment Schemes, *NIST Special Publication*, 800, 56C, 2018.

[3] Boudia, O. R. M.; Senouci, S. M.; Feham, M. (2017). Elliptic curve-based secure multidimensional aggregation for smart grid communications, *IEEE Sensors Journal*, 17(23), 7750-7757, 2017.

[4] Cui, J.; Shao, L.; Zhong, H.; Xu, Y.; Liu, L. (2018). Data aggregation with end-to-end confidentiality and integrity for large-scale wireless sensor networks, *Peer-to-Peer Networking and Applications*, 11(5), 1022-1037, 2018.

[5] Elshrkawey, M.; Elsherif, S. M.; Wahed, M. E. (2018). An enhancement approach for reducing the energy consumption in wireless sensor networks, *Journal of King Saud University-Computer and Information Sciences*, 30(2), 259-267, 2018.

[6] Elsherif, S. M.; Elshrkawey, M.; Wahed, M. E. (2018). An efficient secure scheme for data aggregation in wireless sensor networks using the additive property of complex numbers, *Journal of Electronics and Information Technology*, Elsevier, 7(12), 2808-2814, 2018.

[7] Engouang, T. D.; Yun, L. (2013). Aggregate over multi-hop homomorphic encrypted data in wireless sensor networks. In *2013 2nd International Symposium on Instrumentation and Measurement, Sensor Network and Automation (IMSNA)*, IEEE, 248-252, 2013.

[8] Harb, H.; Makhoul, A.; Tawbi, S.; Couturier, R. (2017). Comparison of different data aggregation techniques in distributed sensor networks, *IEEE Access*, 5, 4250-4263, 2017.

[9] Hasan, M. Z.; Al-Rizzo, H.; Al-Turjman, F. (2017). A survey on multipath routing protocols for QoS assurances in real-time wireless multimedia sensor networks. *IEEE Communications Surveys & Tutorials*, 19(3), 1424-1456., 2017.

[10] Khan, T.; Singh, K.; Abdel-Basset, M.; Long, H. V.; Singh, S. P.; Manjul, M. (2019). A novel and comprehensive trust estimation clustering based approach for large scale wireless sensor networks. *IEEE Access*, 7, 58221-58240, 2019.

[11] Li, X.; Chen, D.; Li, C.; Wang, L. (2015). Secure data aggregation with fully homomorphic encryption in large-scale wireless sensor networks. *Sensors*, 15(7), 15952-15973, 2015.

[12] Li, S.; Kim, J. G.; Han, D. H.; Lee, K. S. (2019). A survey of energy-efficient communication protocols with QoS guarantees in wireless multimedia sensor networks. *Sensors*, 19(1), 199, 2019.

[13] Othman, S. B.; Bahattab, A. A.; Trad, A.; Youssef, H. (2015). Confidentiality and integrity for data aggregation in WSN using homomorphic encryption, *Wireless Personal Communications*, 80(2), 867-889, 2015.

[14] Parmar, K.; Jinwala, D. C. (2014). Aggregate MAC based authentication for secure data aggregation in wireless sensor networks, In *International Conference on Intelligent Computing*,Springer, Cham., 475-483, 2014.

[15] Radhappa, H.; Pan, L.; Xi Zheng, J.; Wen, S. (2018). Practical overview of security issues in wireless sensor network applications, *International journal of computers and applications*, 40(4), 202-213, 2018.

[16] Shim, K. A.; Park, C. M. (2014). A secure data aggregation scheme based on appropriate cryptographic primitives in heterogeneous wireless sensor networks, *IEEE transactions on parallel and distributed systems*, 26(8), 2128-2139, 2014.

[17] Sethi, H., Prasad, D., Patel, R. B. (2011). EIRDA: An energy efficient interest based reliable data aggregation protocol for wireless sensor networks. International Journal of Computer Applications, 22(7), 20-25.

[18] Thiagarajan, R. (2020). Energy consumption and network connectivity based on Novel-LEACH-POS protocol networks, *Computer Communications*, 149, 90-98, 2020.

[19] Shen, Y.; Guo, C.; Wang, L. (2020). Improved Security Bounds for Generalized Feistel Networks, *IACR Transactions on Symmetric Cryptology*, 425-457, 2020.

[20] Schneier, B.; Kelsey, J. (1996). Unbalanced Feistel networks and block cipher design. In *International Workshop on Fast Software Encryption*, Springer, Berlin, Heidelberg, 121-144, 1996.

[21] Xiao, S.; Li, B.; Yuan, X. (2015). Maximizing precision for energy-efficient data aggregation in wireless sensor networks with lossy links, *Ad Hoc Networks*, 26, 103-113, 2015.

[22] Zhao, X.; Zhu, J.; Liang, X.; Jiang, S.; Chen, Q. (2016). Lightweight and integrity-protecting oriented data aggregation scheme for wireless sensor networks, *IET Information Security*, 11(2), 82-88, 2016.

[23] Zhong, H.; Shao, L.; Cui, J.; Xu, Y. (2018). An efficient and secure recoverable data aggregation scheme for heterogeneous wireless sensor networks. *Journal of Parallel and Distributed Computing*, 111, 1-12, 2018.

[24] Zhou, Q.; Qin, X.; Liu, G.; Cheng, H.; Zhao, H. (2019). An Efficient Privacy and Integrity Preserving Data Aggregation Scheme for Multiple Applications in Wireless Sensor Networks. In *2019 IEEE International Conference on Smart Internet of Things (SmartIoT)*, IEEE, 291-297, 2019.

[25] Zhou, Q.; Yang, G.; He, L. (2014). An efficient secure data aggregation based on homomorphic primitives in wireless sensor networks, *International Journal of Distributed Sensor Networks*, 10(1), 962925, 2014.

[26] Zhu, Y. H.; Li, E.; Chi, K.; Tian, X. (2018). Designing prefix code to save energy for wirelessly powered wireless sensor networks, *IET Communications*, 12(17), 2137-2144, 2018.

[27] Zhu, L.; Yang, Z.; Li, M.; Liu, D. (2013). An efficient data aggregation protocol concentrated on data integrity in wireless sensor networks, *International Journal of Distributed Sensor Networks*, 9(5), 256852, 2013.

Cite this paper as:

Elshrkawey, M.; Al-Mahdi, H. (2021). SDA-SM: An Efficient Secure Data Aggregation Scheme using Separate MAC across Wireless Sensor Networks, *International Journal of Computers Communications & Control*, 16(2), 3935, 2021.
https://doi.org/10.15837/ijccc.2021.2.3935