Energy-efficient and reliable data collection in wireless sensor networks

Deepa PUNEETH¹, Nishanth JOSHI¹, Pradeep Kumar ATREY²,
Muralidhar KULKARNI¹,*

¹Department of Electronics and Communication Engineering, National Institute of Technology,
Karnataka, Surathkal, India
²Department of Computer Science, University at Albany, State University of New York, Albany,
New York, NY, USA

Received: 07.02.2017 • Accepted/Published Online: 17.08.2017 • Final Version: 26.01.2018

Abstract: Ensuring energy efficiency, data reliability, and security is important in wireless sensor networks (WSNs). A combination of variants from the cryptographic secret sharing technique and the disjoint multipath routing scheme is an effective strategy to address these requirements. Although Shamir’s secret sharing (SSS) provides the desired reliability and information-theoretic security, it is not energy efficient. Alternatively, Shamir’s ramp secret sharing (SRSS) provides energy efficiency and data reliability, but is only computationally secure. We argue that both these approaches may suffer from a compromised node (CN) attack when a minimum number of nodes is compromised. Hence, we propose a new scheme that is energy efficient, provides data reliability, and is secure against CN attacks. The core idea of our scheme is to combine SRSS and a round-reduced AES cipher, which we call “split hop AES (SHAES)”. Both the simulation results and the theoretical analysis are employed to validate the near-sink CN attack, and a secure reliable scheme using SHAES is proposed.

Key words: Wireless sensor networks, secret sharing, AES, semantic security, energy efficiency

1. Introduction

Wireless sensor network (WSN) applications are becoming ubiquitous in this era of highly networked life. With advances in technology, today’s sensor nodes can gather and process more data, leading to the creation of new applications in recent years. To achieve cost effectiveness, sensor nodes are not typically equipped with tamper-proof facilities [1–13]. Due to their hostile and unattended environment, there is a very high chance of nodes being compromised in a WSN. A compromised node (CN) attack is an attack in which an adversary compromises a certain subset of nodes to passively intercept data packets traversing the compromised nodes [8,9].

Through simulations, this paper validates the vulnerability of secret sharing schemes under the relaxation of a secure area around the BS. We propose a new way of combining the energy-efficient Shamir’s ramp secret sharing (SRSS) method with round-reduced AES symmetric encryption, termed as ‘split hop AES (SHAES)’, to address the CN attack problem. This paper theoretically analyzes the energy efficiency and security of the proposed approach, and shows that the proposed combination achieves both semantic security and reliability in an energy-efficient way.

*Correspondence: mkuldce@gmail.com
1.1. Related work

In the literature, there are various contributions towards security in WSNs. [1] proposes an authenticated group key agreement (AGKA) protocol, and demonstrates how it can tackle node replication and Sybil attacks. In our proposed method, we concentrate on reliability and security by using encryption techniques. [2] discusses efficient and secure routing protocol based on encryption and authentication. This method involves encryption of all communicated packets. In our proposed method, the data are divided into shares and routed via different paths. At the BS, they converge using Lagrange’s interpolation method. We do not encrypt all the shares; the details of the encryption method can be found in Section 2.1. A few similar security management methods can be found in [3], which is based on trust-based management [4], signcryption [5], key predistribution scheme, and so on.

In [6], reliability is achieved through retransmissions. [7] proposes a routing scheme called reliability and multipath encounter routing (RMER), to achieve reliability and energy efficiency. In our proposed approach, we use Shamir’s threshold scheme to achieve the same. With the proposed approach we can achieve both reliability and security. The details of the method can be found in Section 1.2.

The first major contributions to the secure reliable data collection of sensor networks started with H-SPREAD in [8], which used Samir’s secret sharing (SSS) scheme to generate multiple shares of the data. Additionally, a hybrid multipath scheme was used to route the shares. However, the achieved security was low, because fixed multipaths were used to send the data. In addition, the presence of an adversary near the BS was not considered in their approach. The work in [9] addresses the shortcomings of H-SPREAD by using the randomized and highly dispersive nature of routing. This approach increased security compared to H-SPREAD with the help of random dispersion. Network lifetime and security were jointly considered in [10] with a combination of randomized and deterministic multipath routing; however, the approach was very specific to one particular type of deployment strategy, i.e. nodes (circular) and the BS (center of the network area) deployment.

The objective of [11] was to achieve fault tolerance with the help of the secure and efficient disjoint multipath routing strategy. The authors used data duplication and the information dispersal algorithm (IDA) to create multiple data for routing. Similarly, this work assumed the perimeter area around the BS to always be secure. None of the previous approaches considered the possibility of adversaries being near the BS, which is the prime location for obtaining maximum information from the complete network area. If an adversary compromises enough nodes to obtain the threshold shares, then security is lost.

The work in [15] addresses securing the data from aggregator node compromise, making use of secret sharing and multipath routing. In [16], the challenges in the security design of sensor networks are explored. The notion of semantic security in the area of sensor networks was used in their secure network encryption protocol (SNEP) design. Additionally, the authors emphasized that their first choice was the use of the AES block cipher algorithm; however, due to constraints in sensor node memory at that time, they opted for the RC5 algorithm. The need for semantic security in sensor networks to avoid information leakage via eavesdropping has been reported in [17]. The work in [18] considers an ideal linear multisecret sharing (SRSS) scheme in order to provide secure and energy efficient group communications in wireless mesh networks. This approach enhanced energy efficiency, but could not overcome the CN attack problem.

A brief consolidated comparison of the previous works that relate to our proposed core objective is presented in Table 1.
The structure of the paper is as follows: Section 1.2 presents an overview of SSS, Section 1.4 discusses the near sink CN attack, and Section 1.5 discusses the SHAES scheme.

Section 2 presents the details of the proposed work, which explains a combination of the SRSS and SHAES schemes, security analysis, energy efficiency, and reliability analysis. Section 2.4 concludes the paper.

### 1.2. Overview of secret sharing schemes

Shamir’s (t, n) threshold secret sharing (SSS) scheme splits the secret data into n shares [7]. Of these n shares, only t shares are required to reconstruct the complete original data. Any t – 1 shares reveal no information about the message. This desirable property of the SSS scheme helps to generate the data redundancy required to achieve reliability, while still providing information-theoretic security. Share generation is quite simple and is obtained by evaluating the polynomial of degree t – 1 under a Galois field (GF), given by Eq. (1) [20].

\[
S = \left( a_0 + \sum_{i=1}^{t-1} a_i x^i \right),
\]

where \( a_0 \) & are the secret data.

\( a_i \) & are the random data.

\( S \) & is the shares generated for each value of \( x \).

Reconstruction of shares can be achieved by Lagrange’s interpolation method as explained by [20], and is achieved at the BS, which is not usually constrained by resources. To achieve information-theoretic security, the data coefficients are used only at the \( a_0 \) position. This leads to an increase in the required communications to convey the overall data. To reduce the overall communications and, therefore, reduce excessive energy drain, SRSS \((t_1, t_2, n)\), as explained in [20], can be used in WSNs. SRSS allows more data to be used in the polynomial...
computations of S given by Eq. (1), by replacing t with t₂. The first t₀ values of Eq. (1) are obtained from the secret data, and the remaining t₁ (t₁ = (t₂ - t₀)) values are obtained from the random data. Thus, no information about the message is leaked until an adversary can access the t₁ shares. If an adversary has t₁ + 1 shares or more, then information leakage begins and increases until it reaches t₂ shares, where the complete information is obtained [20].

1.3. Assumptions

Sensor nodes are not equipped with tamper-proof facilities. They are prone to be compromised by an adversary, and can perform SRSS and SHAES operations having unique 128-bit keys. The BS is always secure with unlimited energy and processing power, and has complete knowledge of the unique 128-bit key associated with each node. Compared to previous related works, the assumption of the secure area around the BS is relaxed in our approach.

1.4. Near-sink CN attack

The security of SSS lies in the divergence of the shares from the adversary. If the threshold shares converge near the adversary, then secret sharing schemes cannot provide any security. In sensor networks, all the data need to be collected at the BS; therefore, all shares need to converge at the BS. If the adversary compromises a few nodes near the BS, then it can attempt to obtain the shares required to reconstruct the complete data. We term this type of attack as a near-sink CN attack. The near-sink CN attack can compromise the security achieved by SSSs used in WSN applications. To validate the near-sink CN attack, two deployment strategies are considered: centralized sink deployment, where the BS is located at the center of the random nodes, and corner sink deployment, where the BS is located in the corner of the network area containing randomly deployed nodes. A single node is assumed to be compromised by the adversary near the BS. To study the effect of a near-sink CN attack with respect to its distance from the BS, the compromised node is located at different distances from the BS. The transmission power level and communication channel path loss model determine the range and successful reception of data. The near-sink CN attack is tested with different transmission power levels: 0, −1, −3, and −5 dBm. The channel is characterized by the log-normal path loss model, having a path loss exponent of η = 2.4, with slow fading characterized by different standard deviation σ values of 0, 1, 3, and 5 in db. Simulations are carried out in Castalia 3.2, a discrete network simulator with nodes near the BS, which communicates the shares to the BS and has a receiver sensitivity of −95 dBm. Castalia is an open network simulator meant for WSNs and body area networks based on the OMNeT ++ platform. Our simulations were conducted on a network area of 100 × 100 m with 100 randomly deployed nodes. Figure 1a clarifies that under the centralized BS scenario, the circumference area around the BS with a radius of 20 m can overhear 80% of the shares under all transmission power levels. With a transmission power level of 0 dBm, a compromised node can overhear 75% of shares in the perimeter area around the BS with a radius of 45 m. Therefore, a near-sink CN attack is a prominent attack in the applications of WSNs that use secret sharing schemes. When the BS is deployed to any of its network corners, the area around the BS is smaller. Therefore, shares converge well before they reach the BS. Figure 1b justifies that at the 0 dBm power level, a compromised node that is 50 m away from the BS can overhear as much as 90% of shares received by the BS. Although the security achieved by the SSS is higher than the SRSS, energy efficiency is not achieved. Thus, we can conclude that SSSs alone cannot provide adequate security. One way to address this attack is to encrypt the shares. Therefore, in this work we consider reduced AES symmetric encryption, termed as SHAES, to encrypt the shares.
1.5. Split hop AES scheme

The SHAES scheme is a reduced round version of AES with 3 full rounds and with the mix columns (MC) omitted in the last round. The initial AddRoundKey (AK) is retained and acts as the key whitening. From the key schedule, we require only 5 round keys to be generated for use in the AK step. We use a separate name for this round-reduced version of AES: SHAES. Since sensor networks use multihop communication, ‘hop’ is used. Rather than using complete AES, we split it into reduced rounds of 4; hence the term ‘split’ is used. This encryption is used in a somewhat different manner than the normal AES and in combination with secret sharing. The consolidated cryptanalysis attack and its complexity (presented in the literature of 4 round AES) is presented in tabular form in Table 2.

Table 2. Consolidated cryptanalysis attack complexity on four-round AES.

<table>
<thead>
<tr>
<th>Work</th>
<th>Adversary type</th>
<th>Attack type</th>
<th>Data complexity</th>
<th>Time complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>[19]</td>
<td>Highly resource bounded in data</td>
<td>Diff and MiTm</td>
<td>2 CP</td>
<td>$2^{30}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 CP</td>
<td>$2^{32}$</td>
</tr>
<tr>
<td></td>
<td>Highly resource bounded in data</td>
<td>Differential</td>
<td>12 CP</td>
<td>$2^{55}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30 CP</td>
<td>$2^{54}$</td>
</tr>
<tr>
<td></td>
<td>Moderately resource bounded in data</td>
<td>Differential</td>
<td>$2^{11}$ CP</td>
<td>$2^{24}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$2^{14.4}$ CP</td>
<td>$2^{91}$</td>
</tr>
<tr>
<td>[21]</td>
<td>Moderately resource bounded in data</td>
<td>Square</td>
<td>$2^9$</td>
<td>$2^8$</td>
</tr>
<tr>
<td>[22]</td>
<td>Moderately resource bounded in data</td>
<td>Square</td>
<td>$2^9$ CP</td>
<td>$2^9$</td>
</tr>
</tbody>
</table>

2. Proposed work

2.1. SRSS and SHAES combination

The combination of SRSS with encryption is one possible way to overcome the near-sink CN attack. Properly grouping shares and then encrypting them can provide both reliability and security. Figure 2 shows the accurate way of combining the SRSS scheme with the SHAES encryption. Sensor nodes can select proper values for $t_1$, $t_2$, and $n$ based on the application requirements, and can generate the shares. The shares are grouped based on the share numbers, as shown in Figure 2. Only $n - t_1$ grouped shares are encrypted. Since an adversary with only $t_1$ unencrypted shares cannot learn any information about the secret data, we encrypt the remaining shares [20]. The combination scheme provides a resilience of $n - t_2$ share losses. Since only $t_2$ shares are

![Figure 1. Near-sink CN attack under different power levels: a) centralized sink deployment, b) corner sink deployment.](image-url)
required to reconstruct the transmitted data, the BS performs the operations shown in Figure 3 to obtain the original data. If the shares are encrypted, then they are decrypted, and, finally, using interpolation, the data are recovered. All the operations are performed under GF \((2^8)\).

![Figure 2. Graphical representation of achieving the secret sharing and SHAES combination.](image)

### 2.2. SRSS + SHAES security analysis

The definition for probabilistic public-key encryption is provided in [20]. If we extend the same definition to probabilistic symmetric encryption, it would be as follows. Probabilistic symmetric encryption can be defined as six tuples \((P, C, K, E, D, R)\), where \(P\) is the plaintext set space, \(C\) is the set of cipher text space, \(K\) represents the set of key space, \(R\) is the set of randomizer space, and for each key \(k \in K\), \(e_k \in E\) is the encryption rule and \(d_k \in D\) is the decryption rule. The following properties should be satisfied:

1. For each \(e_k: (P, R) \rightarrow C\) and \(d_k : C \rightarrow P\) are functions such that \(d_k (e_k (p, r)) = p\) for every plaintext \(p \in P\) and \(r \in R\). This implies that \(e_k (p, r) \neq e_k (p_1, r)\), if \(p \neq p_1\).

2. For any fixed \(k \in K\) and for any \(p \in P\), define a probability distribution \(f_{(k, p)}(y)\) on \(C\), where \(f_{(k, p)}(y)\) denotes the probability that \(y\) is the cipher text, given that \(k\) is the key and \(p\) is the plaintext (probability
Property 2 states that cipher texts encrypting any two plaintexts should be indistinguishable in polynomial time. This is a desired feature for any security system and provides strong semantic security or message indistinguishability (often used interchangeably). Any block cipher permutation function needs to be bijection (i.e. one-to-one and onto). AES is a block cipher, and for any $k \in K$, $e_k()$, the encryption function is also bijection. Therefore, Property 1 is satisfied. The SRSS + SHAES combination achieves semantic security and comes under probabilistic symmetric encryption. If the adversary using a CN attack or a near-sink CN attack overhears the transmitted message, then they cannot learn any information from the $t_1$ unencrypted shares, and, therefore, cannot reconstruct the original data. The adversary needs to successfully decrypt the encrypted shares to know the information about the data. Since the SRSS + SHAES combination provides semantic security, the adversary does not learn any new information from the encrypted shares and, therefore, will be unsuccessful in decrypting the shares. Thus, the proposed optimized SRSS + SHAES combination overcomes the CN attack problem in WSNs.

Figure 3. Graphical representation of retrieving the original message from the secret sharing and SHAES combination at the BS.
2.3. Energy efficiency and reliability analysis of SRSS + SHAES

Data redundancy is necessary to achieve reliability. Therefore, as the data size and redundancy increases, communication energy drain increases. The objective is to achieve reliability with a minimum increase in data size. Figure 4a shows the percentage increase in data size for 1024 bits of data using SRSS and SSS schemes in order to meet various reliability requirements indicated by \( n - t_2 \). For instance, to afford to lose 7 shares (i.e. \( n - t_2 = 7 \)) while the SRSS scheme with \( t_0 = 8 \) presents a 100% increase in data size, the SSS scheme with \( t_0 = 1 \) would have an 800% increase in data size.

**Figure 4.** a) Communication and reliability analysis of SRSS + SHAES. b) Computational overhead analysis in terms of number of multiplication operations for different approaches.

Note that SRSS with \( t_0 = 1 \) is the same as the SSS scheme. Furthermore, reliability requirements depend on the wireless channel properties and, therefore, vary based on the channel conditions. Computation energy depends on the complexity of the schemes, hardware implementation, and the processor. One general way to analyze computation overhead is to analyze the complexity of the schemes by calculating the number of complex operations. Major complex operations involved in SRSS and SHAES are addition and multiplication operations under GF \( (2^8) \). Among these multiplication operations are the most computationally expensive operations [20]. In AES, the MC involves multiplication operations. Addition operations are involved in AK, MC, and key schedule. The number of addition and multiplication operations required to realize AES is calculated to be 752 and 576, respectively. Referring to AES, the number of addition and multiplication operations needed to realize SHAES is equal to 272 and 192, respectively. Furthermore, the polynomial evaluation of SRSSs involves addition, multiplication, and exponential operations. Figure 4b shows the total number of multiplication operations needed to realize the combination of SSS with AES and SHAES, encrypting each group of shares. The exponential operations can be realized with repeated multiplication operations. The number of multiplication and addition operations needed to communicate the data size of \( D \) bytes using SRSS are given by Eqs. (2) and (3), respectively. Figures 4b and 5a show the multiplication operations needed to communicate the data size of 1024 bits using the SSS and SRSS schemes, respectively. For instance, at SSS \( (t_2 = 5, n = 5) \) and SSS \( (t_2 = 10, n = 10) \), the numbers of multiplication operations needed are 6400 and 57,600, respectively. At SRSS \( (t_0 = 4, t_2 = 5, n = 5) \) and SRSS \( (t_0 = 8, t_2 = 10, n = 10) \), the number of multiplication operations needed are 1600 and 7200, respectively. As \( t_2 \) increases, the number of multiplication operations increases, as indicated in Figures 4b and 5a. Therefore, to reduce the computation burden under secret-sharing
schemes, one needs to select parameters with a lower $t_2$, resulting in fewer multiplication operations.

\[ \frac{D}{t_0} \times n \times \sum_{i=1}^{t_2} i\# \]  

(2)

\[ \frac{D}{t_0} \times n \times (t_2-1)\# \]  

(3)

The proposed optimized SRSS and SHAES combination encrypts a minimum group of shares (total size = 128 bits), as explained in Figure 2. If the encrypting share group size is less than 128 bits, then extra bits must be added. As operations are performed in GF ($2^8$), each share will consist of 8 bits. A group of 16 shares would be equal to the required 128 bits. Therefore, a proper selection of SRSS parameters, based on the data size lengths, helps to reduce the extra bits and achieves better communication energy efficiency. Figure 5b shows the number of polynomial evaluations for different data sizes using various SRSS parameters. The SRSS parameter combinations for different data sizes in the lower half of the demarcated line of Figure 5b are not efficient, as they require extra bits to make the share group size reach 128 bits. Since SRSS analysis in this paper is restricted up to $t_0 = 8$, the lowest data size that results in an efficient combination using $t_0 = 8$ is equal to 1024 bits. Therefore, the results of 1024-bit data size are presented in this paper. The number of multiplication operations needed to realize the combination of SRSS with AES and SHAES encrypting each group of shares are shown in Figures 5c and 5d, respectively. For instance, at SRSS $[(t_0 = 4, t_2 = 5, n = 5) \text{ and } (t_0 = 8, t_2 = 10, n = 10)]$, the number of multiplication operations needed for SRSS + SHAES and SRSS + AES are [3520 and 9120] and [7360 and 12,960], respectively. The combination of SSS + SHAES with minimum share group encryption has high computational and communication overheads. The proposed optimized combination of SRSS + SHAES achieves low computational and communication overheads, as explained in Figures 6a and 6b. For instance, at SRSS $[(t_0 = 4, t_2 = 5, n = 5) \text{ and } (t_0 = 8, t_2 = 10, n = 10)]$ the number of multiplication operations needed for optimized SRSS + SHAES is [3136 and 8736]. From these results, we can conclude that the proposed optimized SRSS + SHAES with minimum share group encryption is highly energy-efficient compared to other combinations. Computational overhead is analyzed using the numbers of multiplication operations. The different approaches are given different levels of computational overhead, ranging from lowest to highest, based on the overall trend observed with the increase in $t_2$ and data size. Consolidated analysis of different objectives is presented in tabular form in Table 3.

<table>
<thead>
<tr>
<th>Works</th>
<th>Objective</th>
<th>CN attack</th>
<th>Computation overhead in terms of (x) operations</th>
<th>Communication overhead in terms of data size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous works [8–10,18]</td>
<td>Only SSS</td>
<td>Yes</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Only SRSS</td>
<td></td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>Different approaches examined</td>
<td>SSS + AES</td>
<td>No</td>
<td>Highest</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>SSS + SHAES</td>
<td>No</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>SRSS + AES</td>
<td>No</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>SRSS + SHAES</td>
<td>No</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>SSS + AES optimized</td>
<td>No</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>SRSS + SHAES optimized</td>
<td>No</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>
3. Conclusion

In this paper, the vulnerability of secret sharing schemes under the relaxation of a completely secure area around the BS to near-sink CN attacks was validated through simulation results. Simulations were carried out using MATLAB and Castalia, a discrete network simulator. The theoretical analysis validated the achieved energy
efficiency and desired semantic security. The proposed combination works independently from the underlying routing schemes. Therefore, it can be easily incorporated into existing related works in secure data collection of WSNs.

The following could be considered for future work in the area. Proposing an energy-efficient data reconstruction scheme for secret sharing optimized selection on the ramp secret sharing parameters, based on the sensor network routing constraints and cryptanalysis of the proposed probabilistic symmetric encryption scheme.

References


