A DNA-BASED PRIVACY-PRESERVING SCHEME IN SMART-GRID

Woud M. Abed
Department of Basic Science, College of Dentistry
University of Baghdad, Baghdad, Iraq

ABSTRACT

Smart grid utility provider collects consumers’ power consumption data for three main reasons: billing, analysis, and operation. Billing needs coarse-grained data where there are no, or minimal, privacy concerns. While analysis and operation needs fine-grained data which can highly explore consumers’ privacy. Hence, consumers might be reluctant to allow for operational metering to protect their privacy. This paper presents detail description of a reliable DNA-based privacy-preserving (DNAPP) scheme in smart grid. DNAPP assures robust authentication, confidentiality, message integrity, and non-repudiation across the smart grid as well as assuring high consumers’ privacy. The scheme demonstrates many good security features, such as: high complexity of $O(n!)$, light-weight, scalable, minimum overhead, no cryptography key exchange between the communicating parties as each of them can determine the key locally and independently. This scheme does not require any level of modifications to the existing smart grid infrastructure or smart meter. It only requires some software modifications.

KEYWORDS

Smart grids; smart meters; privacy-preserving, security services; DNA; random permutation.

1. INTRODUCTION

The legacy power grid has several limitations resulting in misuse and mismanagement of power resources and supply. Fortunately, the tremendous development in information and communication technology (ICT) as well as smart devices technologies empower the development of traditional power grid and introduce the concept of smart grid (SG). Hence, SG is considered as the integration of the latest ICT solutions with the traditional power grid in order to optimize power management. SG links the different grid actors from the consumers’ smart meters (SM) to the end-head system, such as the SG authority server (SGAS) [1]. Introducing new technologies into SG raises new issues, mainly privacy and security, such as SM readings may disclose private information about customers’ daily life and habits [2-4].

The core component of a smart grid is the SMs, which are intelligent electronic devices that collect and record information on power consumption, and communicate that information to the utility through communication channels for monitoring, billing, analysis for advanced power demand and generation management. SMs also enable consumers to remotely control the operation of their network connected devices to reduce power consumption through the application software of the SM owner (SMO) [5].

Communicating sensitive consumers’ power consumption and control data between SM and SGAS from one side and between SM and SMO from the other side risks consumers privacy and security and must be carefully addressed. Furthermore, it threatens consumers’ privacy by disclosing fine-grained consumption data and consumer’s power usage behavior. In particular, as data travels through several networks, secure end-to-end communication based on strong authentication and encryption mechanisms are crucial to assure privacy-preserving, data confidentiality, and integrity of exchanged data[5, 6].

DOI:10.5121/ijcis.2019.9301
Many solutions have been developed to address the problem of security and privacy in SGs, such as implementation of strong encryption and authentication techniques [4-6]. Deoxy ribo Nucleic Acid (DNA) cryptography is emerging as a new promising cryptographic field where DNA is used to carry the information or to be used as an alternative data encoding approach. During the last two decades, many DNA-based algorithms have been developed and used for data cryptography and cryptographic key generation [7, 8].

In this paper, we develop a new DNA-based privacy-preserving (DNAPP) scheme that ensure a secure data exchange between the main components of the SG shown in Fig. (3) (e.g., SGAS, SM, and SMO), authentication, message integrity, non-repudiation, as well as maintain consumers’ privacy. The scheme identifies two types of sessions, one between SGAS and SM, the other between SM and SMO. So that each SM locally determines two cryptographic keys one for each session, SGAS locally determine a cryptographic key for each SM it communicates with, and SMO determines its relative cryptographic key. Hence, in DNAPP, SMs’ reading, control data, or aggregation data can be exchanged without being disclosed to any unauthorized users.

This paper is divided into five sections. This section introduces the main theme of this paper. Section 2 provides a brief background on the concept of SG, SG network model, and major security requirements of SGs. A review of some of the most recent and related research is presented in Section 3. The new DNA-based authentication scheme is discussed in Section 4. In Section 5, conclusions are drawn and recommendations for future research are pointed out.

2. SMART GRID ARCHITECTURE AND NETWORK MODEL

In this section, we provide a description of SG architecture and smart grid network model. In addition, this section defines the main security requirement of SGs [2-6].

2.1. Smart Grid Architecture

The architecture of a typical SG network is depicted in Figure 1, which shows that SG consists of three main components; these are: smart meters, concentrator, and central management facilities. Smart meters usually installed within the consumers’ premises and they frequently transmit their readings to the local electricity supplier. Each smart meter communicates with other meters in the neighbourhood and with the local concentrator through a wire or wireless mesh network. The concentrator acts as the local provisional data collection unit, which communicates with the central management facilities such as the electric utility and grid operator through wired/wireless communication channel, and reliably reports the aggregated meter readings in the neighbourhood to the facilities.

2.2. Network Model

There are different models for the SG communication infrastructure. One of the most appropriate models is the wireless-wired multi-layer architecture. In this model, the SMs in the neighbourhood are communicating with the collector through wirelessly, on the other hand, the collector communicates with central management unit of the grid operator through wireless/wired communication channels. In case, the collector may not be able to directly connect with every SM in the neighbourhood; then the unconnected SM should establish at least one communication path through a set of other intermediate SMs in range to the collector, as shown in Figure 2. SMs becomes a critical component of smart home, which is used to remotely control the operation of home appliances to reduce electricity bill at consumers’ premises.
Routing the data through intermediate SMs means that these meters may have knowledge to portion or all of consumers’ consumption data that are of interest of attackers and becomes targets of various types of attacks. Without protection, it is easy for an attacker to passively spy on the data at a certain SM or actively forge the aggregation value through compromised SMs. Therefore, it is important to protect the confidentiality and authenticity of the aggregation data in SGs [3].

2.3. Threat Model and Security Requirements

In SG networks, we consider both active attacks (e.g., data manipulation or fake data injection) and passive attacks (e.g., eavesdropping) by adversaries. To alleviate the effects of these attacks, the following security requirements are necessary to ensure secure smart metering in SGs [9]:

1. Data confidentiality (consumer’s privacy): The power consumption and any other data exchanged through the SM is considered as private information of the consumer; therefore, it should not be disclosed to the concentrator or other consumers on the routing path during data aggregation and delivery. Also, they may be exposed to passive attacks, such as eavesdropping. Thus, they should be secured against such attacks.
2. Data integrity: Dishonest or compromised SMs in the network could manipulate the intermediate metering data during aggregation, causing inaccurate aggregation results. Thus, manipulation of the aggregate by active inside attack from the compromised meters should be detected by the concentrator.

3. Sender authentication: Defending against any active outside attacks, such as false data injection attack by outside adversaries, the concentrator should be able to ensure the authenticity of the SMs’ identities on the routing path that have contributed to the data aggregation.

3. LITERATURE REVIEW

This section is divided into two subsections; the first one reviews some of the most recent and related researches and surveys on privacy-preserving in SG communications. The second discusses DNA-based cryptography, and how DNA-based methods are used for assuring security requirements of authentication and confidentiality.

3.1. Privacy-Preserving

Koo et. al [3] investigated the current status of security and privacy in SG, in particular the secure aggregation and authentication of metering data in future SG. Hur et. al [5] and Li et. al [6] developed a distributed incremental data aggregation solutions, in which data aggregation is processed at all SMs along data route from the source meter to the collector. They used homomorphic encryption to secure the data along its route and also protect user privacy. The solutions are particularly suitable for SGs with repetitive data aggregation tasks.

A security analysis of a SM authentication protocol was also described and investigated by Uto et. al [10]. They investigated the critical vulnerabilities found in the authentication process and show how it can be failed by a brute-force attack in few hours. Nabeel et. al [11] developed a strong hardware-based SM authentication protocol using a physically unclonable function (PUF) technology. It provides a reliable authentication of SMs and efficient key management to ensure the integrity and confidentiality of communicating data between SMs and the utility provider. One advantage of this protocol is that it does not require modifications to the existing meter communication and it only requires some software update.

For interesting researchers, there are a number of surveys that have been published in the last few years discussing the requirements, challenges, issues, and solutions for future privacy-preserving in SGs. In what follows, we shall mention some of these surveys.

Kumar et al. [12] presented a brief overview of real cyber-attack targeting smart metering network. They classified these threats into three categories: threats in system-level security, threats and/or theft of services, threats to privacy. Based on the category of threats, they derived a set of security and privacy requirements for SG metering networks. Furthermore, they discussed and identified the pros and cons of each of the various schemes that have been proposed to address these threats. Finally, they investigated the open research issues to explore future research directions in SG networks.

Ferrag et al. [2] presented a comprehensive survey of SG privacy-preserving schemes. In particular, they examined a number of privacy preserving schemes published between 2013 and 2016 for SGs privacy preserving. Based on the outcomes of their survey, they pointed out several recommendations for further research.
Liu et al. [13] surveyed cyber-security and privacy issues in SG. They find that privacy in SGs may be addressed by proper adaptation of existing information and network security technologies. Yan et al. [9] discussed specific SG security requirements along with challenges and available solutions. They described several solutions that were implemented or tested on real environments for privacy protection, integrity, authentication, and trusted computing. They briefly described seven encryption and anonymization techniques; however they did not provide any comparison between them. Another survey similar to [9] presented by Wang et al. in [14] in which they devoted a section for privacy, where they presented and analyzed some cryptographic, authentication and key management schemes, along with case studies. They concluded that a tradeoff between latency and privacy is a major smart grid security concern.

Komninos et al. [15] discussed open issues, challenges and countermeasures for SG and smart home security. They also discussed several privacy-preserving techniques that are based on anonymization, encryption, perturbation, verifiable computation models and obfuscation. They recognized the need of a legal framework specific to privacy-preserving in the SG, the establishment of new key management techniques and new aggregation mechanisms. Another survey by Tan et al. [16] discussed and investigated data generation, data acquisition, data storage, data processing, and data analytics aspects of SG security. Many other surveys are available in the literature focus on privacy-preserving for SGs [17-19].

### 3.2. DNA-Based Cryptography

DNA cryptography is a promising research approach that emerged with the evolution of DNA computing field. Several DNA-based algorithms have been developed and used in many applications, such as data encryption, private key generation, authentication, etc. [7, 20]. Gupta and Jain [21] developed a method for image encryption based on DNA computation technology, where they first, generate a secret key using a DNA sequence and modular arithmetic operations. Then each image pixel undergoes encryption process using the key and DNA computation methods. Zhang et al. [22] also developed image encryption algorithm based on DNA sequence addition operation.

Varma and Raju [23] analysed the performance of different DNA-based matrix manipulation and secret key generation schemes. Liu et al [24] presented an encryption method using DNA complementary rule where piecewise linear chaotic map is used for permutation and then substitution is performed using complementary rule. Rakshit et al. [25] developed a DNA-based cryptography method. The theoretical analysis and implementations show that the method is efficient in terms of computational speed, storage requirement, and transmission. Also it is very powerful against certain attacks. Najaftorkaman and Kazazi [26] developed and examined the performance of a novel DNA-based cryptography, where they use DNA coding to convert binary data to DNA strings.

All of the aforementioned research papers and surveys that are related to privacy-preservation for SGs do not mention the use of the promising DNA cryptography in SG privacy-preserving. Therefore, in this paper, we propose a new scheme that uses the DNA concept to ensure privacy-preserving in SG.

### 4. DNA-BASE PRIVACY-PRESERVING (DNAPP) SCHEME

This section provides a detail description of the DNAPP scheme, which provides authentication, data confidentiality, message integrity, and non-repudiation security services in smart grid networks. In SG, each smart meter (SM) is assumed to be interconnected with the SG authority server (SGAS) from one side and with the smart meter owner (SMO) on the other side, whether this interconnection is direct (one-hop) or indirect through intermediate nodes (multi-hops). The communications among these three objects (e.g., SGAS, SM, and SMO) is illustrated in Figure (3).
The main parameters that are identified the objects in Figure 3 are given below:

1. SGAS ID (SID). It is a unique number that identified a particular SGAS.
2. SM Serial No.(SN). It is a unique number that identified a particular SM.
3. Consumer ID (CID). It is a unique ID issued by the SGAS specific to a certain SM.
4. SGAS Secret ($K_{SS}$). It is issued by SGAS and it can be changed according to the network and operation environment. Each SGAS↔SM channel has its own $K_{SS}$.
5. SMO Secret ($K_{OS}$). It is issued by the SMO, and it be changed according to the network and operation environment. Each SM↔SMO channel has its own $K_{OS}$.

All of the above parameters are agreed on between the communicating objects during the installation stage. SGAS knows the first four parameters only (SID, SN, CID, $K_{SS}$), while the SMO knows all of them except the fourth one (SID, SN, CID, $K_{OS}$). The SM knows all the parameters, and it uses the first four to determine the cryptography key for communication with SGAS ($K_{S}$), and uses the first three plus the fifth to determine the cryptography key for communication with SMO ($K_{O}$).

The calculation module of the DNAPP scheme consists of three main steps:

1. Step 1: SGAS, SM, and SMO use the related SID, SN, CID, $K_{SS}$, and $K_{OS}$ to locally determine the DNA-based cryptography key($K_{s}$) as shall be described below. In fact, SM determines two keys one to communicate with SGAS ($K_{S}$), and the other to communicate with SMO ($K_{O}$).

   $$K_s = f(SID, SN, CID, K_{SS})$$

   $$K_o = f(SID, SN, CID, K_{OS})$$ (1) (2)

2. Step 2: The sending party(SGAS, SM, or SMO) then uses the relative pre-determined key (either $K_{S}$ or $K_{O}$) to encrypt the communicating data using one of the standard symmetric ciphering techniques such as RC4 [27]. The two parties must agree on the ciphering technique that will be used in a session, and the parties may agree to use different encryption techniques for each session or task depending on the network environment and the criticality of the communicating data. This mathematically can be expressed as:
\[ C_{SGAS\rightarrow SM} = E_{K_S}(M_{SGAS\rightarrow SM}) \] (3)
\[ C_{SM\rightarrow SMO} = E_{K_O}(M_{SM\rightarrow SMO}) \] (4)

Where \( M_{SGAS\rightarrow SM} \) and \( C_{SGAS\rightarrow SM} \) are the plain and encrypted message to be exchanged between SGAS and SM, \( M_{SM\rightarrow SMO} \) and \( C_{SM\rightarrow SMO} \) are the plain and encrypted message to be exchanged between SM and SMO, \( E \) is the encryption algorithm, and \( K_S \) and \( K_O \) are as defined above.

3. Step 3: The receiving party (SGAS, SM, or SMO), then uses the relative predetermined key (either \( K_S \) or \( K_O \)) to decrypt the communicating data using the pre-agreed encryption algorithm. This mathematically can be expressed as:
\[ M_{SGAS\rightarrow SM} = D_{K_S}(C_{SGAS\rightarrow SM}) \] (5)
\[ M_{SM\rightarrow SMO} = D_{K_O}(C_{SM\rightarrow SMO}) \] (6)

Where \( D \) is the decryption algorithm. Other parameters are as defined above.

4.1 DNA-Based Cryptography Key

In this work, the DNA-based cryptography keys are determined as follows:

1. Determine two random permutations \( P \) each of size \( n \) using any public seed pseudorandom permutation algorithm (also called key-based pseudorandom permutation algorithm), such as the algorithms in [28, 29]. In this case, we determine two permutations, one for SGAS-SM channel \( (P_S) \) and the other for SM-SMO channel \( (P_O) \). The seed for the permutation is the hash of the string produced from concatenated the relative parameters, such that:
\[ P_S = f(\text{Hash}(\text{SID} \& \text{SN} \& \text{CID} \& \text{SS})) \] (7)
\[ P_O = f(\text{Hash}(\text{SID} \& \text{SN} \& \text{CID} \& \text{SO})) \] (8)

The MD5 or SHA1 hash functions can be used in the calculation [27].

2. The permutation \( P(P_S \text{or} P_O) \) is used to determine the DNA-based cryptography key \( (K_S \text{ or} K_O) \) as follows:
   a. Encode each element of the permutation \( P \) to their equivalent binary value. The number of bits representing each element \( (m) \) is calculated as: \( m = [\ln(n) / \ln(2)] \). So that the length of the binary sequence \( (L) \) representing the permutation \( (P) \) of length \( n \) is \( L = m \cdot n \) bits.
   b. Convert each two consecutives bits to an integer value \( (00 \rightarrow 0, 01 \rightarrow 1, 10 \rightarrow 2, \text{and} 11 \rightarrow 3) \).
   c. Store these integer values in a vector \( V \) of size \( L \).
   d. Split the vector \( V \) into \( m \) vectors \( (V_1, V_2, V_3, \ldots, V_m) \) each of size \( n \).
   e. Permute the vectors \( (V_1, V_2, V_3, \ldots, V_m) \) using the permutation \( P \) to produce permuted vectors \( (PV_1, PV_2, PV_3, \ldots, PV_m) \).
   f. The \( n \) elements of the DNA key can be calculated as:
\[ DNA(k) = (\sum_{i=1}^{m} PV_i) \mod 4 \] (for \( k \) 1 to \( n \)) (9)
3. Convert each DNA base to its 2-bit equivalent value (A as 00→00, C as 1→01, G as 2→10, and T as 3→11). This will yield the DNA cryptography key.
The steps in 2 and 3 above are equally applied to determine $K_S$ and $K_O$ using $P_S$ and $P_O$, respectively. It can be clearly recognized that the determined DNA components are randomly distributed over the DNA-based generated key without any previous estimation on the occurrence of each DNA component.

The length of the permutation and key can be easily increase to accommodate any number of SMs and also it can be increased to enhance the strength of the authentication secret without overloading SM or SGAS. Furthermore, the secrets agreed between SM and SGAS and SM and SMO make it very difficult for attackers to explore the encryption/decryption key as it can be dynamically changed.

Now we have a key that is uniquely and locally determined at all communicated parties, so that any symmetric data encryption algorithm (e.g., RC4) can be used to cipher the personal data (SM reading, control data, or aggregation data) to be exchanged between the SM and the SGAS. The key needs not to be exchanged between the SM and SGAS as each of them can determine the key independently. The same is applied with SMO as he/she can agree with the SM on a certain secret to generate entirely different key while exchanging data between the SM and the SMO, for example exchanging some control data to switch ON/OFF devices connected to the SM at home or work.

5. CONCLUSIONS

This paper presents detail description of a new DNA-based privacy-preserving (DNAPP) scheme. In this scheme, the SMs’ reading, control data, or aggregation data can be exchanged between the SMs and the SGAS without being disclosed to any intermediate SM or the concentrator(s). The new scheme demonstrate an excellent security performance in terms of high complexity of $O(n!)$ as well as it is light-weight because it is requires a simple computation procedure. Furthermore, it does not require cryptography key exchange between the communicating parties as each of them can calculate the key locally and independently.

The scheme is also scalable to cover any size of SG regardless the number of SMs with minimum overhead. Also, it is very difficult for hackers or adversaries to explore the cryptography/authentication key as it is depend on some key parameters known to the SGAS, SM and SMO as well as some secret parameters that is can be frequently changed and agreed upon between the communicating parties (i.e., the SM and the SGAS or the SM and SMO). Finally, it must be well recognized that this scheme is at its early stage and it needs further tests and evaluate its performance against different security attacks.

REFERENCES


International Journal on Cryptography and Information Security (IJCIS), Vol. 9, No.3, September 2019


AUTHORS

Woud M. Abed is a member of academic staff at the Department of Basic Sciences, College of Dentistry, University of Baghdad (Baghdad, Iraq). She received her B.Sc degree in Computer Science from the Department of Computer Science, Al-Rafidain University College (Baghdad, Iraq) in 2003, and her M.Sc degree in Computer Networks, Informatics Institute for Higher Studies, University of Technology (Baghdad, Iraq) in 2005. Her research interests include: robotic, genetic algorithms, cryptography and steganography, image processing, and computer security.