A Secure Cloud Computing Based Framework for Big Data Information Management of Smart Grid

Devadharshini L, Nivetha R.
Department of Computer Science and Engineering,
PITS.

***Abstract*:- Smart grid is a technological innovation that improves efficiency, reliability, economics, and sustainability of electricity services. It plays a crucial role in modern energy infrastructure. The main challenges of smart grids, however, are how to manage different types of front-end intelligent devices such as power assets and smart meters efficiently; and how to process a huge amount of data received from these devices. Cloud computing, a technology that provides computational resources on demands, is a good candidate to address these challenges since it has several good properties such as energy saving, cost saving, agility, scalability, and flexibility. In this paper, we propose a secure cloud computing based framework for big data information management in smart grids, which we call "Smart-Frame." The main idea of our framework is to build a hierarchical structure of cloud computing centers to provide different types of computing services for information management and big data analysis. In addition to this structural framework, we present a security solution based on identity-based encryption, signature and proxy re-encryption to address critical security issues of the proposed framework.**

I. INTRODUCTION

1.1 Big Data Analysis in Smart Grid

SMART grids have recently been adopted in electronic grid renovation plans of many countries, replacing traditional power grids. One of the reasons is that compared to traditional power grids, smart grids bring significant improvement in the efficiency, reliability, economics, and substantiality of electricity services [18]. As an example, the ENEL Telegestore project in Italy [38], which is widely regarded as the first com-mercial project using smart grid technology, delivers annual savings of approximately 500 million Euros [37]. Following the success of Telegestore, several other smart grid projects have been proposed. They include the Hydro One project [10] in Canada, the Evora InovGrid project [26] in Portugal, and the Modellstadt Mannheim (Moma) project [36] in Germany. While smart grids bring in several benefits to electrical power grids, their deployment is often limited to small regions (e.g., within a city or a small province).

There areseveral challenges that prevent smart grids to be deployed at a larger scale (e.g., in the whole country), one of which is infor-mation management that is related to information gathering, information storing, and information processing [5], [14], [17]. Since there are a large number of front-end intelligent devices, managing a huge amount of information received from these devices is not an easy task. In a preliminary estimation at one utility, the amount of data required to process transac-tions of two million customers could reach 22 gigabytes [44] per day. It is definitely a big challenge to manage this set of big data, which may include the selection, deployment, moni-toring, and analysis of smart grid data. More importantly, a real-time information processing is usually required in the smart grid. Any delay may cause a serious consequence in the whole system which has to be avoided as much as possible.

1.2 Assistance from Cloud Computing

Cloud computing has become popular recently due to several advantages over traditional computing models. Typical advantages include flexibility, scalability, agility, energy efficiency, and cost saving [24]. For this reason, it has been expected to be a dominant computing model in the future. By employing cloud computing in smart grids, we not only address the issue of large information management but also provide a high energy and cost saving platform. It is because 1) the framework can scale very fast to deal with changes in the amount of processing information and 2) it can provide a high utilization of computing resources.

Actually, prior to our work, initial efforts have been devoted to prove that cloud computing can satisfy requirements of information management in these systems [4], [40]. In particular, in [40], properties of smart grid and cloud com-puting were analyzed to prove the relationship between them. Furthermore, in [4], use cases of a smart grid were dis-cussed to understand detailed requirements of information management, and cloud computing properties were studied to show that they meet the requirements. Nevertheless, none of these works comes up with a concrete design for information management in smart grids besides rather abstract analyses.

1.3 Our Approach

Motivated by the previous work, in this paper, we introduce a design of Smart-Frame, a flexible, scalable, and secure information management framework for smart grids based on cloud computing technology. Our basic idea is to build the framework at three hierarchical levels: top, regional, and end-user levels in which the first two levels consist of cloud computing centers while the last level contains end-user smart devices. The top cloud computing center takes responsibility of managing general devices and accumulation of data across the regional cloud computing centers which are placed in the lower level in the hierarchy. The regional cloud computing centers are in turn in charge of managing intelli-gent devices, which have lower hierarchical level than the regional cloud computing centers in specific regions (e.g., within a city), and processing data of these devices.

In addition to this general framework, we propose a security solution for the framework based on identity-based encryption (IBE) and signature [6], [43], and identity-based

proxy re-encryption [22]. Providing information security for smart grids is very important since much of the information in smart grids is sensitive and needs to be strictly protected. Information leakage in smart grids can lead to vulnerabil-ities that affect not only individuals but also the whole nation because leaked information can be used to launch attacks to both individuals and the whole smart (power) grids at the national level.

The main idea of our security solution for the Smart-Frame is to allow all the involved entities, i.e., top and regional cloud computing centers and end-users to be represented by their identities which can be used as encryption keys or signature verification keys. The entities in the lower level can use the identities of higher-level entities to encrypt their data for secure communication with the entities in the higher level. For example, the regional centers use the top cloud's entity to encrypt their messages. By employing an identity-based re-encryption scheme, the information storages, which are com-ponents of regional clouds, can re-encrypt the received confi-dential data from the end-user devices so that services requested by the end-users decrypt and process the confiden-tial data without compromising the information storages' pri-vate keys. One of the obvious benefits we can gain from applying identity-based cryptography to the Smart-Frame is that through using identities rather than digital certificates which depend on traditional public key infrastructure (PKI), we can save significant amount of resources for computation and communications and resolve scalability issues. The saving gained from the elimination of digital certificate in the big data environment is especially momentous.

1.4 Our Contributions

To summarize, our contributions in this paper are twofold:

We introduce Smart-Frame. A cloud computing based framework for big data information management in smart grids, which provides not only flexibility and scalability but also security features.

We present a security solution for the proposed framework based on identity-based encryption and proxy re-encryption schemes, which provides secure

communication services for the Smart-Frame. We further implement the prototype of our proposed solution to show its practicality.

The rest of this paper is organized as follows. In Section 2, we review the related work. Through Sections 3 and 4, we present the Smart-Frame. In particular, Section 3 focuses on the general architecture of the Smart-Frame while Section 4 deals with its security issues and proposed a solution based on identity-based cryptography. Finally, we demonstrate a simple prototype implementation in Section 5 and conclude the paper in Section 6.

I. Related Work

In this section, we review the related work about smart grid information management (Section 2.1), smart grid security management (Section 2.2), and finally the basics of identity-based encryption and proxy re-encryption schemes respec-tively (Sections 2.3 and 2.4).

2.1 Smart Grid Information Management

Smart grid information management usually involves three basic tasks: information gathering, information processing, and information storing. For information gathering, since smart grids have to collect information from heterogeneous devices at different locations, the main research challenge is to build efficient communication architecture. Several solu-tions have been proposed to address this challenge and most of them can be found in the recent surveys such as [17], [46], and [5]. In terms of information processing, data integration also lays a challenge as information can be received from a number of devices, which may use different data structures to handle the information. Fortunately, a proposal for standardization of data structures used in smart grid applications has recently been proposed to address this issue of data inter-operability [25]. However, how to process a large amount of data efficiently still remains as a big challenge. Cloud computing appears to meet this demand and also satisfy challenges of information storing. As a result, initial works on cloud computing and smart grids have been produced. In [40] properties of smart grid and cloud computing were analyzed to prove that cloud computing is a good candidate for information man-agement in smart grids. Similarly, in [4], use cases of a smart grid were discussed to understand detailed requirements of information management and cloud computing properties were studied to show that they meet the requirements. These two works are different from ours in that they only presented analysis while we introduce a concrete design for the platform as well as a security solution for it.

information processing of smart meters while Zhang et al.

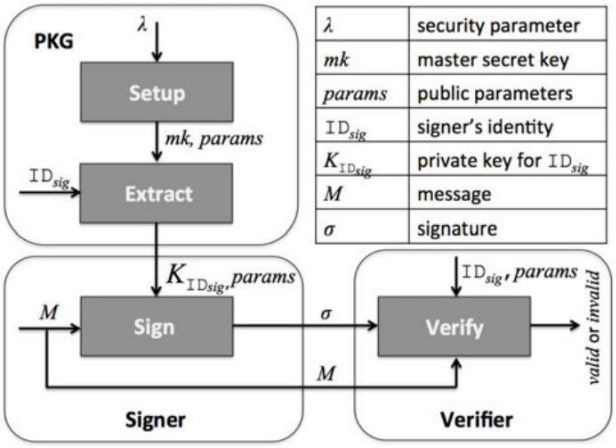

Fig. 2. Overview of identity-based signature.

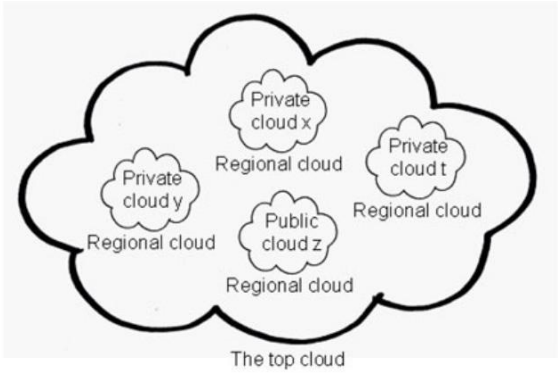

Fig. 3. Overview of the architecture of the Smart-Frame.

identity-based encryption was accomplished by Boneh and Franklin [6] in 2001.

2.4 Identity-Based Proxy Re-Encryption

Proxy re-encryption lets a proxy to transform a ciphertext produced under Alice's public key in such a way that the transformed ciphertext can be decrypted under another party Bob's private key. The concept of proxy re-encryption was first introduced by Mambo and Okamoto [34], whose main goal was to achieve efficiency better than "decrypt-and-encrypt" approaches. The first fully functioning proxy re-encryption scheme was proposed by Ateniese et al. [1]. Compared with the previous approaches, their proxy re-encryption scheme was unidirectional, so it does not require delegators to reveal their secret keys to anyone in order to allow proxy to re-encrypt their ciphertexts.

Since Ateniese et al.'s work, numerous proxy re-encryp-tion schemes with various functionalities have been pro-posed. Among them, the identity-based proxy re-encryption scheme proposed by Green and Ateniese [22] is closely related to our Smart-Frame. In an identity-based proxy re-encryption scheme, a delegator allows a proxy to transform an encryption under Alice's identity into one encrypted one under Bob's identity. The proxy then uses re-encryption keys to conduct the transformation without being able to learn any information about the plaintext. Also, no informa-tion about the private keys of Alice and Bob would be deduced from the re-encryption keys. Note that identity-based proxy re-encryption combine the two functionalities of IBE and proxy re-encryption without compromising the security. Note also that Green and Ateniese's identity-based proxy re-encryption scheme [22] is based on the pairing like Boneh and Fraklin's IBE scheme.

II. SMART-FRAME

In this section, we discuss our proposed Smart-Frame from three main perspectives: system architecture (Section 3.1), logical components (Section 3.2), and information manage-ment (Section 3.3).

3.1 General System Architecture

The overall architecture of the Smart-Frame is shown in Fig. 3. In this architecture, a smart grid can be divided into several regions each of which is managed by a cloud computing center that can be setup from either a public cloud or a private cloud. The role of a regional cloud comput-ing center is to manage intelligent devices in the region as well as to provide an initial processing for information received from these devices. Besides regional cloud comput-ing centers, there is a special cloud computing center at the top level, which is in charge of managing and processing data for the whole grid. In each of these cloud computing centers, the following cloud computing services could be deployed:

Infrastructure-as-a-service (IaaS). This type of service forms the backbone of the system. It helps to provide resources on demand for all applications and serv-ices deployed in the system. Main tasks of informa-tion management in smart grids such as information gathering, information processing, and information storing, are all executed inside this layer of service.

Software-as-a-service (SaaS). While IaaS is the back-bone of the system, all smart grid services will be deployed as SaaS at the top of the system. Examples include services that allow customers to save or optimize their energy usage such as Google Power Meter [21].

Platform-as-a-service (PaaS). PaaS provides tools and libraries to develop cloud computing applications and services. Salesforce [41] is a typical PaaS exam-ple, which provides libraries to develop some spe-cific types of applications in saleforce or fieldforce domains. In smart grid domain, since a number of applications could be required to follow special secu-rity requirements and have to allow lawful intercep-tions, it is useful to have a general PaaS that has already integrated these requirements to implement applications.

Data-as-a-service (DaaS). DaaS could be deployed to provide useful information for statistics purpose. Since smart grid data is often extremely large, it is useful to provide such statistics services for users. Statistics can be used for optimization purposes for not only electricity users but also electricity pro-viders at different levels.

3.2 Logical Component View

Among cloud computing services presented in Section 3.1, while IaaS is the backbone of the system, we classify other services into clusters according to functionality they provide

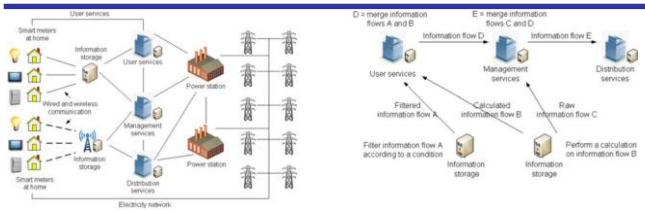

Fig. 4. Functional cloud computing service clusters.

in order to ease the management. In our framework, we propose to use four main functional clusters as follows:

Information storages. These are main storages keeping all smart grid information received from front-end intelligent devices. These storages are designed to accept information from different transportation modes through both wired and wireless channels. For optimization purpose, statistics services are also located in this cluster.

General user services. This type of services consists of all services an electricity user needs to use. Typical examples are services that allow users to monitor, control or optimize the usage of their electric utili-ties. The majority of SaaS fall into this type of service. PaaS that provides libraries for user services also falls into this cluster.

Control and management services. This type of services includes all services needed for system management such as governance service, monitoring service, task scheduling service, and security service.

Electricity distribution services. This type of services is directly related to electricity distribution. Examples

are distribution management service, optimization service, and quality of service measurement.

The above four types of services are illustrated as in Fig. 4. Note that besides information storages, all other types of services can be linked to the electricity grid. Note also that among these functional clusters, while information storages and user services usually exist in regional clouds, management and distribution services can be found in both regional and top clouds.

3.3 Information Flow Management

Since smart grids need to handle huge amount of data, it is extremely important to manage information flows effi-ciently. In the Smart-Frame, we suggest a centralized service to manage information flows. This service takes inputs as both information requests from service clusters and general statistics (e.g., the amount of information, time of arrival) from information storages. Using these inputs, the service generates an information flow schedule, which specifies sources and destinations of information flows as well as how they are processed (e.g., which specific operators are applied on information flows and where they are applied).

Both information storages and services clusters need to follow this schedule for execution. Fig. 5 shows an example of information flow schedule. Note that since information amount and requests in smart grid may change with time, each information flow schedule has an expiry time. After this time, a new schedule has to be generated and distrib-uted again to all parties.

III. Security Solution for Smart-Frame

4.1 Security of Smart-Frame Supported by Identity-Based Cryptography

Since security is a major concern in the smart grids, it is of great importance for our Smart-Frame to provide a solution to address that. As mentioned earlier, one of the huddles for widely deploying security solutions based on public key cryptography is the high cost for maintaining PKI. We envision that Identity-based cryptography can be a good solution (though it is not perfect) for resolving this problem since identity-based cryptography has the following advantages in regards to the Smart-Frame security.

1) Under traditional public key cryptography, each participating entity must locate and verify the public keys of the receivers. This is especially burdensome for end-user devices in our Smart-Frame, which are usually assumed as limited in power and network-ing capacity.

2) Although traditional public key cryptography is scalable in theory, a number of issues regarding user interfaces for maintaining public-key certificates (involving certificate revocation) have to be resolved. However, since any identifier strings can serve as encryption key or signature verification keys, identity-based cryptography could provide better scalability for the system. This is important in Smart-Frame in which numerous end-user devices can join and leave the system often.

3) the users' identities and master key of the private key generator, no secure database can be required for the system based on identity-based cryptography.the data which they not entitled to process. An example is illustrated in Fig. 7 where the end-user agrees to let serv-ices A, B, and C to receive and use its data.

The details of our security framework are more formally described as follows. Note that TC, RC, and EU denote top cloud, regional cloud, and end-user respectively. We assume that the regional cloud consists of information storage IS and services $SerA_1$; $SerA_2$; . . . ; $SerA_n$. For the sake of convenience of the description, we assume that SerA repre-sents each service, i.e., RC ¼ fIS; SerAg. We also assume that IS and SerA are independently run, and do not share confi-dential information nor collude. Additionally, we assume that the basic identity-based cryptographic schemes IBE and IBS, which the following security framework unitize, are as described in Section 2.3.

Key Generation

- Setup: Given a security parameter , the PKG generates a secret master key mk and a set of parameters params. The PKG distributes params to all the clouds and end-users.

- ExtractTCKey: Upon receiving a top cloud's identity TC, the PKG generates a private key K_{TC} associated with TC by running the private
  - key extraction algorithm Extract providing TC as input. We denote this process by K_{TC}
  - ExtractRCKeyðparams; mk; TCÞ.
- ExtractISKey: Upon receiving an identity of the information storage in the regional cloud,
  - denoted by IS, the PKG generates a private key K_{IS} associated with IS by running the private
  - key extraction algorithm Extract providing IS as input. We denote this process by K_{IS}
  - ExtractISKeyðparams; mk; ISÞ.
- ExtractServiceKey: Upon receiving an identity of
  - the service A in the regional cloud, denoted by SerA, the PKG generates a private key K_{SerA} asso-ciated with SerA by running the private key
  - extraction algorithm Extract providing SerA as input. We denote this process by K_{SerA}
  - ExtractServiceKeyðparams; mk; SerAÞ.
- ExtractEUKey: Upon receiving an end-user's identity EU, the PKG generates a private key K_{EU} associated with EU by running the private key
  - extraction algorithm Extract providing EU as input. We denote this process by K_{EU} ExtractEUKeyðparams; mk; EUÞ.
  - Encryption to Information Storage (in Regional Cloud)

- Encrypt2IS: Each end-user can encrypt a mes-sage M into a ciphertext C_{IS} by running the IBE encryption algorithm Encrypt with params and the identity IS of the information storage
- in the regional cloud. We denote this process by C_{IS} Encrypt2ISðparams; IS; MÞ.
- DecryptIS: Each regional cloud can decrypt a received ciphertext C to M by running the
- IBE decryption algorithm Decrypt with the pri-vate key K_{IS} associated with the information
- storage's identity IS. We denote this process by
- M DecryptISðparams; K_{IS}; C_{IS}Þ.
- Encryption to Top Cloud
- Encrypt2TC: Each information storage in the
- regional cloud can encrypt a message M into a ciphertext C_{TC} by running the IBE encryption algorithm Encrypt with params and the top cloud's identity TC. We denote this process by
- C_{TC} Encrypt2TCðparams; TC; MÞ.
- DecryptTC: The top cloud can decrypt a received ciphertext C to M by running the IBE decryption
- algorithm Decrypt with the private key K_{TC} asso-ciated with the top cloud's identity TC. We denote this process by M DecryptTCðparams; K_{TC}; C_{TC}Þ.
- Proxy Re-encryption by Information Storage

- RKGen: Providing its own private key K_{IS}, its identity IS and the server A's identity SerA as
- input, the information storage in the regional cloud generates a re-encryption key $RK_{IS!SerA}$.
- We denote this process by $RK_{IS!SerA}$ RKGen ðK$_{IS}$; IS; SerAÞ.
- – Reencrypt: The information storage in the regional cloud re-encrypts the ciphertext C_{IS} using the re-encryption key $RK_{IS!SerA}$ and obtains a ciphertext C_{SerA}. We denote this process
- by C_{SerA} ReencryptðRK$_{IS!SerA}$; C_{IS}Þ.
- DecryptService: The service A decrypts C_{SerA}
- using its private key K_{SerA}. We denote this by
  - DecryptServiceðK$_{SerA}$; C_{SerA}Þ.
- Signature Generation by End-Users
- SignEU: Each end-user can generate a signature s for a message M using the private key K_{EU}
- associated with its identity EU. We denote this process by s SignEUðparams; K_{EU}; MÞ.
- VerifyEU: Any party can verify a signature s for some message M using params and the iden-tity of the end-user, EU. We denote this process
- by d VerifyEUðparams; EU; s; MÞ, where d is either "accept" or "reject".
- Signature Generation by Entities in Regional Cloud
- SignIS: Each information storage in the regional cloud can generate a signature s for a message
  - using the private key K_{IS} associated with its
- identity IS. We denote this process by s SignISðparams; K_{IS}; MÞ. Each service in the regional cloud (denoted by SerA as a representative) can also generate a signature i the same way.

IV. Prototype Implementation

5.1 Overview

As a proof-of-concept, we implemented a simple prototype for our proposed framework. In our implementation, all cloud computing centers, both regional centers and the top cloud center, were built based on Eucalyptus [16], a popular open source cloud computing platform. By using Eucalyptus, we aim to provide infrastructure-as-a-service to the platform users. We chose Eucalyptus for our framework because of the following reasons.

It is fully compatible with the industry standard amazon web services cloud APIs.

It supports all major virtualization technologies including Xen, KVM, and VMware vSphere.

It can be developed and extended easily and be installed smoothly on all major Linux OS distribu-tions such as Ubuntu, RHEL/CentOS, openSUSE, and Debian.

On top of the Eucalyptus platform, to support the secu-rity for the framework, we provide the following services.

Identity registration. Identity registration is used to register identities of all components that need to send or

receive information in the framework. As an example, smart meters, intelligent sensors and all other front-end devices need to register their identi-ties before they are allowed to send information to the cloud storage. On the other hand, cloud comput-ing components and services need to register their identities before they can receive or provide informa-tion. When an identity is registered, a private key associated with the identity is generated for the reg-istered component.

Data encryption and data decryption. Data encryption is used to encrypt data before it is sent through the net-work. In general, before sending the data, the sender uses the identity of the target receiver as the key to encrypt the data. Given that the target receiver is the only one who holds the private key to decrypt the data, this way the security of data is retained. On the other hand, data decryption is used by a receiver of ciphertext to decrypt the previously encrypted data(ciphertext) to obtain original data.

These above services were implemented based on the Java-based cryptographic library for pairing operation called JPair developed by Dong [13]. Given the platform and basic security services, all information management tasks as well other types of services can be implemented on top of the platform.

5.2 A Specific Scenario of the Platform Usage

As an example, we give a specific scenario of the platform usage. Suppose that a regional center (server) identified by a string "AD_EC" for controlling electricity in Abu Dhabi, United Arab Emirates. Suppose also that a smart meter identified by a string "SM1" which resides in a household in AlMata, a region of Abu Dhabi, will encrypt a message regarding the daily usage of electricity. Let this message be "SM1||50kW||AlMata". This scenario is realized in our proposed security framework .

In the first step, the regional center (server)'s identity AD_EC is registered using the identity registration ser-vice. After registration, the private key associated with the identity is issued. Fig. 8a demonstrates this process which is displayed in console. Note that AD_EC.sk denotes the private key associated with regional server (center)'s identity AD_EC.

In the next step, the smart meter uses the regional center (server)'s identity AD_EC to encrypt its mes-sage regarding the daily usage of electricity by call-ing the data encryption service. Fig. 8b demonstrates this process, displayed in console. Note that the ciphertext, consisted of three components U, V and W, is longer than the original message due to the redun-dancy added by the encryption algorithm to guaran-tee strong security.

In the final step, the regional center uses its private key AD_EC.sk to decrypt the ciphertext by calling the data decryption service. Fig. 8c demonstrates this process, displayed in console. The decrypted message is "SM1||50kW||AlMata" which is interpreted as "The daily usage of electricity recorded in Smart Meter 1 in household in AlMata is 50kW".

V. CONCLUSION

In this paper, we have introduced the Smart-Frame, a gen-eral framework for big data information management in smart grids based on cloud computing technology. Our basic idea is to set up cloud computing centers at three hier-archical levels to manage information: top, regional, and end-user levels. While each regional cloud center is in charge of processing and managing regional data, the top cloud level provides a global view of the framework. Addi-tionally, in order to support security for the framework, we have presented a solution based on identity-based cryptog-raphy and identity-based proxy re-encryption. As a result, our proposed framework achieves not only scalability and flexibility but also security features. We have implemented a proof-of-concept for our framework with a simple
identity-based management for data confidentiality. Our immediate next step is to also support proxy re-encryption for the framework.

REFERENCES

[1] G. Ateniese, K. Fu, M. Green, and S. Hohenberger, "Improved proxy re-encryption schemes with applications to secure distrib-uted storage," ACM Trans. Inf. Syst. Security, vol. 9, no. 1, pp. 1–30, 2006.

[2] J. Baek, Q. Vu, A. Jones, S. Al-Mulla, and C. Yeun, "Smart-frame: A flexible, scalable, and secure information management frame-work for smart grids," in Proc. IEEE Int. Conf. Internet Technol. Secured Trans., 2012, pp. 668–673.

[3] A. Bartoli, J. Hernandez-Serrano, M. Soriano, and M. Dohler, "Secure lossless aggregation for smart grid M2M networks," in Proc. IEEE Conf. Smart Grid Commun., 2010, pp. 333–338.

[4] K. P. Birman, L. Ganesh, and R. V. Renesse, "Running smart grid control software on cloud computing architectures," in Proc. Work-shop Comput. Needs Next Generation Electric Grid, 2011, pp. 1–33.

[5] Z. Bojkovic and B. Bakmaz, "Smart grid communications architec-ture: A survey and challenges," in Proc. 11th Int. Conf. Appl. Com-put. Appl. Comput. Sci., 2012, pp. 83–89.

[6] D. Boneh and M. K. Franklin, "Identity-based encryption from the weil pairing," in Proc. 21st Annu. Int. Cryptol. Conf. Adv. Cryptol., 2001, vol. 2139, pp. 213–229.

[7] X. Boyen, "A tapestry of identity-based encryption: practical frameworks compared," Int. J. Appl. Cryptograph., vol. 1, no. 1,
pp. 3–21, 2008.

[8] C.-K. Chu, J. K. Liu, J. W. Wong, Y. Zhao, and J. Zhou, "Privacy-preserving smart metering with regional statistics and personal enquiry services," in Proc. 8th ACM SIGSAC Symp. Inf., Comput. Commun. Soc., 2013, pp. 369–380.

[9] G. M. Coates, K. M. Hopkinson, S. R. Graham, and S. H. Kurkow-ski, "A trust system architecture for SCADA network security," IEEE Trans. Power Delivery, vol. 25, no. 1, pp. 158–169, Jan. 2010.