Wireless Pers Commun (2013) 73:23–50
DOI 10.1007/s11277-012-0939-1

Towards Privacy Protection in Smart Grid

Sherali Zeadally · Al-Sakib Khan Pathan · Cristina Alcaraz · Mohamad Badra

Published online: 16 December 2012

Abstract The smart grid is an electronically controlled electrical grid that connects power generation, transmission, distribution, and consumers using information communication technologies. One of the key characteristics of the smart grid is its support for bi-directional information flow between the consumer of electricity and the utility provider. This two-way interaction allows electricity to be generated in real-time based on consumers' demands and power requests. As a result, consumer privacy becomes an important concern when collecting energy usage data with the deployment and adoption of smart grid technologies. To protect such sensitive information it is imperative that privacy protection mechanisms be used to protect the privacy of smart grid users. We present an analysis of recently proposed smart grid privacy solutions and identify their strengths and weaknesses in terms of their implementation complexity, efficiency, robustness, and simplicity.

Keywords Authentication · Confidentiality · Energy · Privacy · Smart grid

S. Zeadally (✉)
University of the District of Columbia, Washington, DC, USA
e-mail: szeadally@udc.edu

A.-S. K. Pathan
Department of Computer Science, International Islamic University Malaysia (IIUM), Kuala Lumpur Malaysia
e-mail: sakib.pathan@gmail.com; sakib@iium.edu.my

C. Alcaraz
University of Malaga, Malaga Spain
e-mail: alcaraz@lcc.uma.es

M. Badra
College of Applied Sciences - Sohar, Ministry of Higher Education, Sohar Oman
e-mail: mbadra@gmail.com

1 Introduction

1.1 Traditional Electric Power Grid

The traditional electric power grid is typically seen as a transmission system that transfers electricity from bulk generation systems (e.g., nuclear systems, hydroelectric systems, wind farms, and others) to power distribution substations (as shown in Fig. 1), and each substation finally delivers electricity at a low voltage to their end users. The energy production and distribution schema are supervised by a centralized control system, known as Supervisory Control and Data Acquisition (SCADA) systems, in charge of mapping and visualizing any operational activity in the field as well as controlling the storage and demand of power. In fact, SCADA systems can remotely and locally control the power transmission and distribution based on the current demand and peak loads thereby minimizing unnecessary power generation.

Nonetheless, the architecture illustrated in Fig. 1, composed of an important set of interconnected engineering resources, has significantly evolved in recent decades with the integration of new Internet Protocol (IP)-based technologies. Network convergence technologies [1] have opened up control connections of the electrical power grid to external networks through the Transmission Control Protocol/Internet Protocol (TCP/IP).

Over the last few years, we have witnessed important advances in hardware, software, and communication technologies that have resulted in the widespread deployment of Information and Communication Technologies (ICTs), smart and mobile devices, software applications and architectures [2,3].

1.2 The Smart Grid

The advent of ubiquitous computing and communication technologies have also led to a major shift toward a smarter, interactive, and dynamic electric grid that is viewed as the next generation of the 21st century electrical grid, widely known as the smart grid (as shown in Fig. 2). The smart grid provides significant benefits in terms of its support for bi-directional flow of information both to the appliances and devices inside the customer premise and back to the utility provider using IP-based communications.

According to the conceptual model of the National Institute of Standards and Technology (NIST), a smart grid is a complex infrastructure based on a set of seven chief domains [4]: bulk generation, energy distribution, power transmission, operation and control, market, service

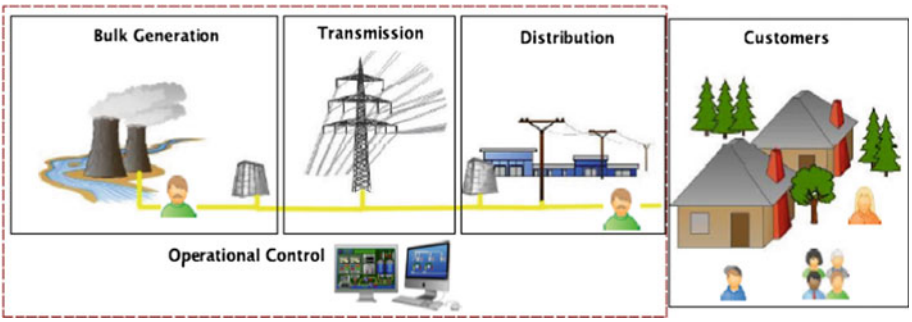

Fig. 1 Architecture of the traditional electric power grid

providers, and customers. Each domain comprises heterogeneous elements that include organizations, buildings, individuals, systems, system resources and other entities. The backhaul communication and the Internet are crucial for connecting the different entities involved such as customers and utility systems through an Advanced Metering Infrastructure (AMI) [5]. An AMI is an interface with the capability for managing and interacting with smart meters and utility business systems through a bi-directional communication. This communication tries to substitute the one-way Advanced Meter Reading (AMR) approach by enabling business utilities or providers to notify their customers of electricity pricing at any time, providing them with customizable services to manage their power consumption themselves in addition to controlling the demand in real time. The smart meter is defined as an advanced meter (usually an electrical meter, but could also integrate or work together with gas, water, and heat meters) that measures energy consumption in much more detail than a conventional meter does. Future smart meters are envisaged to communicate information back to the local utility company for monitoring voltage loads and for billing purposes.

There are several technologies and applications that have been integrated to perform as one in an AMI system [6] including: smart meters, wide-area communications infrastructure, Home (local) Area Networks (HANs), Meter Data Management Systems (MDMS), and operational gateways working as main collectors. AMI are solid state programmable devices that can perform many functions allowing users to perform intended tasks by inputting a sequence of instructions into its processing unit and memory. Among some of the tasks that a smart meter can do are [6]: time-based pricing, collecting consumption data for consumer and utility, net metering, loss of power (and restoration) notification, better access and data to manage energy, decision and selection of rate options, remote turn on/turn off operations, load limiting for "bad pay" or demand response purposes, energy prepayment, power quality monitoring, meter tampering and energy theft detection, costs reduction in wrong estimations of billings, service and operational reduction in traditional tasks of metering reading, or communications with other intelligent devices or appliance devices in the home. Although all these tasks may not be supported by a particular meter and there might be other tasks that it can do, the overall idea is that smart meters make it possible to add some kind of "*intelligence*" to the network and individual features of each residential consumer. The main parameters for managing the demand side are not the hourly consumption of energy, but the maximum demand. Each application area will have different needs, because their customers also will have different demands. The privacy related issue here is that for proper functioning of AMI system, very detailed and often precise information about user's electricity usage is needed. Hence, while this smart system could offer many great benefits, it takes away significantly from the level of privacy a user may like to have.

1.3 Hardware and Software Technologies used in Smart Grid

The smart meter is the core element on the customer side of the AMI system. A smart meter is usually an electrical meter that records consumption of electric energy in intervals of an hour or less and communicates that information at least daily back to the control utility for monitoring and billing purposes [7]. Figure 2 (left hand side figure) shows a smart meter and an old style meter, whereas Fig. 2 (right hand side figure) illustrates the general hardware architecture of these devices. In particular, smart meters are typically based on microcontrollers and support optimal digital signal processing functions for power quality measurement features, a shared memory (RAM, ROM and flash), communication ports (e.g., USB, Ethernet, optical Ethernet, serial connector) with communication capabilities such as Universal Asynchronous Receiver/Transmitter (UART), RS-485, Wi-Fi and ZigBee.

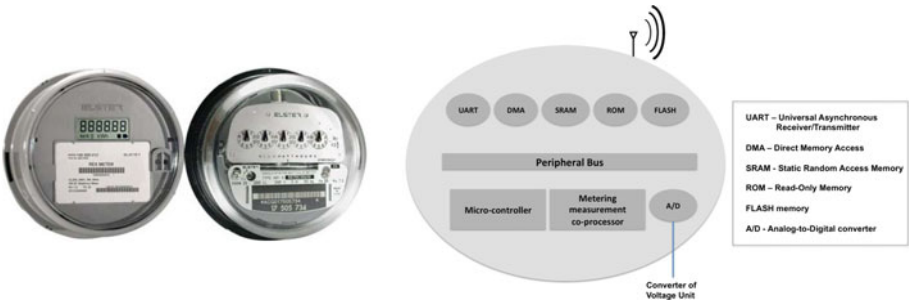

Fig. 2 A modern solid state smart meter (*left*) and an older electromechanical watt hour meter (*right*) [6]

For energy load and distribution within a smart grid, the system can also be subdivided into small smart microgrids interconnected through communication infrastructures in charge of sending commands (i.e., actions), alarms (i.e., information about the current states of the infrastructure), and readings (i.e., measurements of a context such as temperature or voltage) to control systems (i.e., the SCADA Central system). A microgrid basically consists of a localized system of electricity generation, energy storage and load of power resources which are normally connected to a traditional centralized system. This means that a microgrid can also function autonomously (i.e., in island mode). As shown in Fig. 3 the microgrids may be connected to energy production systems (e.g., renewable, renewable non-variable, non-renewable/non-variable systems [4]). Any information (e.g., alarms, commands, readings of voltage values) has to be forwarded to the control system by means of a wide variety of technologies and IP-based protocols. Some of these technologies and protocols are shown in Table 1.

1.4 Contributions of This Paper

The deployment and adoption of smart grid technologies have opened up several security issues at the levels of the consumer, the communication, and of the energy provider. Security

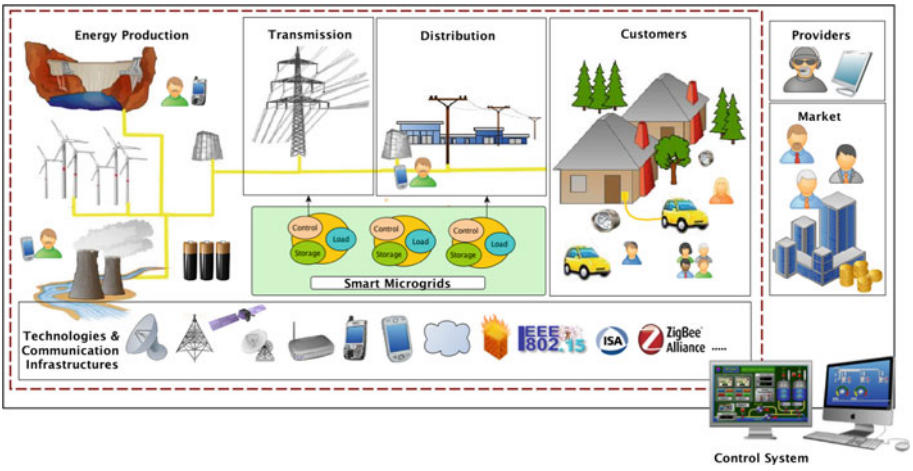

Fig. 3 Architecture of the smart grid

Table 1 Some technologies and IP-based protocols used in smart grid

Power generation, transmission and distribution systems	Power control systems	AMI	HAN
Technologies			
Mobile cellular technology (e.g., Third Generation/Fourth Generation (3G/4G), Universal Mobile Telecommunication System (UMTS), General Packet Radio Service (GPRS), Global System for Mobile Communications (GSM)), satellite, WiMAX, Mobile Broadband Wireless Access (MBWA), microwaves systems, optical fiber, bluetooth,Wi-Fi, Wireless Sensor Networks (WSNs), Ethernet, and others			
IP-based Protocols			
IEC 61850, IEC 61968, IEEE 1815 (DNP3), Modbus, IEC 60870, IEC 61400-25, proprietary, and others	Zigbee, WirelessHART, ISA100.11a, IEC 61850, IEC 62351, IEEE 1815 (DNP3), Modbus, IEC 60870, CIM, ICCP/TASE 2.0, proprietary, and others	ZigBee Smart Profile, ANSI C12.2, and others	ZigBee Smart Profile, Z-Wave, ANSI C12.2, IEEE 1547, HomePlug, and others

aspects such as confidentiality, authentication, authorization, integrity, and non-repudiation for smart grid technologies are currently being extensively investigated and various innovative solutions are being proposed in the literature. The authors of [8] provided some of the early insights into how to smarten electricity systems leaving out security related issues. Lu et al. [9] reviewed the security threats towards communication networks in the smart grid ecosystem and evaluated the impact of these threats. Steven et al. [10] focused on smart grid security areas such as trust, communication, and device security. McDaniel et al. [11] discussed several issues resulting from the deployment of the smart grid infrastructure and presented various security and privacy challenges in the smart grid. In their work, McDaniel et al. distinguish security and privacy issues. They argue that security solutions defend against various forms of frauds and attacks on the system while privacy solutions make data inaccessible to unauthorized parties. Although the work of McDaniel et al. provided a very limited contribution to the issue of smart grid privacy, it did highlight its importance in future smart grid deployment and adoption. Since then, the study of privacy in smart grid has started to generate a lot of interest in the research community and industry particularly when it comes to the collection and the use of energy consumption data collected from homes that are using the smart grid technology.

In contrast to most previous works on smart grid which have focused mostly on smart grid security issues, the primary goal of this work is to review, discuss, and analyze recent smart grid privacy solutions that have been proposed in the literature and identify their strengths and weaknesses. In doing so, it is our hope that our findings will help designers and implementers to develop and implement cost-effective, efficient privacy solutions for the smart grid.

The rest of this paper is organized as follows. In Sect. 2 we highlight several privacy issues with smart grid deployments. Section 3 describes design architectures and approaches that have been recently proposed to protect the privacy of smart grid users. Section 4 presents current laws and regulations that can be used to partially protect, to some extent, the privacy of smart grid users. Finally, our concluding remarks are presented in Sect. 5.

2 Privacy Issues with Smart Grid

2.1 Basic Privacy Concepts

Privacy may be defined as the claim of individuals, groups or institutions to determine when, how and to what extent information about themselves is communicated to others [12]. The notion of privacy may vary from person to person, and from culture to culture. It could also be defined as the right to informational self-determination, i.e., individuals must be able to determine for themselves when, how, to what extent and for what purpose information about them is communicated to others [56]. This term is often related to an entity's (individual, group, or institution) identity or anonymity. As human beings, each of us likes to keep some information about ourselves confidential while we like to express some information to draw a distinct line with others or to make a presence in the society that we live in. Similarly, a group or institution may have some information for disclosure to the public while sensitive information must be protected from being disclosed to unwanted parties. The unwanted parties may include individuals who are not the members of the group or institution, other groups or institutions, a person with short-term membership, or a deliberate intruder (attacker) attempting to retrieve information illegitimately.

The definition and boundaries of privacy tend to vary among different societies and cultures and as such, there is no clear list of categories of privacy that can be applicable for all. However, four major types of privacy are generally recognized:

- **Personal Privacy:** this includes mainly body privacy and territorial privacy. Body privacy varies among individuals in terms of the types of clothing one wears to protect the body. Territorial privacy means making a boundary or to create a barrier between the person and others. This can be implemented by erecting walls/ fences/screens, by using cathedral glass/partitions, by maintaining a distance, besides other ways.
- **Information Privacy:** this kind of privacy is mainly related to passing of information over various media and could also be called communications privacy. Some of the notable information privacies are:
  - *Internet privacy*: the ability to determine the kind of information one reveals or withholds about oneself over the Internet, who has access to such information, and for what purposes one's information may or may not be used.
  - *Financial information privacy*: information about own bank account, amount of money, transaction details, debt, etc.
  - *Medical privacy*: information about a person's health conditions.
  - *Political privacy*: political stance such as who a person may have voted for.

  Information privacy also means how someone expresses matters about himself/herself in any field. People are sometimes willing to give up information about themselves not because they are ignorant or because they are being tricked by evil corporations, but because it can sometimes be in their best interests to do so [13,14]. Such information can be posted on the Internet or via social networks or other channels the person is involved with. So, in such a case, a person may judge the benefit of exposing such information which he/she may like others to know but not through himself/herself directly, may be to avoid the accountability or responsibility of such apparent "*leak*" of information.
- **Organization Privacy:** this includes the confidential information about an organization such as business strategies, loss and profit statistics, current trend in the market, future products, potential customers, transaction details, and similar information. An organi-

zation may put some information in the public arena for transparency (which will show the ethical standard of the organization, commonly accessible by anybody) and declares certain information as classified, which is a categorization applied to information that a government or a group claims as sensitive. Prominent examples of organizational security could be often associated with trade secrets and national security.

- **Spiritual and Intellectual Privacy:** this kind of privacy includes a person's spiritual nature, of his feelings and his intellect. A person may have certain religious beliefs but he may not like to express it to others. It may be because of the adverse or hostile environment. Also, a highly intelligent person may act as dumb or may not like to show his intelligence in all gatherings. For example, a person working in a research group may restrain from showing all his talents to others so that others may not take his ideas away without giving proper credit or it may be that the person is selfish or he may like not to actually get involved in intellectual contribution in the group for some personal reasons.

As the meanings of privacy are different in various scenarios, there are other ways of looking at it. Pedersen [15,16] described six types of privacies related to a man's personality: (i) solitude, (ii) isolation, (iii) anonymity, (iv) reserve, (v) intimacy with friends, and (vi) intimacy with family. Solitude is the most complete state of privacy that individuals can achieve. It is a type of privacy in which the individual is alone and unobserved. Pedersen differentiates between isolation termed as alone and away from others and solitude defined as alone by oneself and free from observation by others. Anonymity is a type of privacy that occurs when it is possible to move around in public or for example, browsing through the Internet without being recognized or being the subject of attention. Reserved behavior includes examples of low self disclosure. Finally, any kind of intimacy is a type of privacy that relates to an individual's or group's desire to promote close personal relationships. All of these personal traits of human beings need to be studied and thoroughly understood while making any policy related to privacy in any sector, because the same human beings are the beneficiaries or users of these systems.

2.2 The Need for Privacy in Smart Grid

In a smart grid network, key questions regarding setting the policies on user data privacy are [17]: Who owns the data of the customer? How is the access to and use of customer data regulated? Who guarantees privacy and security of customer data (e.g., against risk of surveillance or criminal activity)? Will sale or transfer of customer data be allowed, and under what terms and to whose benefit? In jurisdictions with retail choice, are measures needed to ensure competing electricity providers have access to customer data on the same terms as the incumbent utility?

In fact, rival electricity providers may compete to dominate the market, and their access to users' electricity usage pattern and behavioral information could be very crucial. The electricity providers or provider agents may use the user data to determine their business strategies and special packages/offers. In an open market environment, such data could be partially collected after the offers are made public and some information is available for all, but if privacy is breached beforehand and specific user data is available to some parties, then these electricity providers may have unfair gains. Appropriate privacy policies may restrict or mitigate or resolve such use of unfair means in setting business strategies. All these issues explain why the privacy of data of smart grid users is a very critical issue both for users and the electricity providers.

The privacy of smart grid users is a very important issue. The strong integration of Information and Communication Technologies (ICTs) for the smart grid's operation introduces different types of privacy concerns. Depending on the method how the consumer (or, user) uses electricity and recharges it, the privacy of the user can be affected by two usage scenarios namely:

- ***The user recharges electricity balance via personal interaction (private mode)***: for instance, the user goes *in person* to the electricity provider's agent and recharges his "smart-electricity-card" similar to a credit/debit card that can be reloaded and placed into the electricity meter. The other personal interaction may happen via the phone or in person by going to the agent and getting a new recharge/reload number similar to that used in many places for pre-paid mobile phone balance/validity extension. The customer can also obtain a recharging number obtained from a pre-paid card. This method does not reveal the identity of the person who has purchased the card which is later used in the electrical meter to do the re-loading task. It is worth pointing out that the authorization number will need to be validated and authenticated before electricity consumption. When this number is entered from any home or building (connected to smart grid), it passes through an authentication process during which information could be stored by the utility company or one of its designated agents. This information needs privacy protection measures in place.
- ***The user recharges electricity balance via the Internet (public mode):*** if any website or online system is used and the balances are adjusted via payment through some bank account or other payment methods, then all the cybersecurity-related privacy issues must be considered. When a web interface is used and there is a back-end database, web attacks (such as Structured Query Language (SQL) injection [18]) could affect the privacy of the user by disclosing not-to-be-exposed data from the back-end database. The web-based (i.e., online) form to recharge the user's electricity balance could be made as simple as requiring a single identification number from the user. The privacy issue in this process is whether the user wants to be known at the time of recharging a balance for future electricity usage. In fact, user's information can be used by different departments/branches of the electricity provider. The user may choose who could access the information and who could not. An instance of personal preference can be the option of receiving company related news, updates or offers of newly introduced packages or benefits from the electricity supplier company to the user's email address. For managing user's own preferences, agent technology [19] could be used, in which each subscriber/user is assigned an agent representing the user's interests. Each service can also be assigned an agent to reap the most benefit. A service agent could negotiate with subscriber agents about information and authorizations versus the quality of the offered service.

The level of personal information involved and used will dramatically increase with the modernization of the grid. Smart meters and smart appliances could lead to a data explosion of intimate details of daily life. However, at this point, it is quite unclear as to who will gain access to this information besides the customer's utility provider and control utilities. With the deployment of the smart grid, energy measurements can take place at much shorter intervals (unlike at the end of the billing cycle as in conventional methods).

Currently, there are several types of concerns related to the privacy and security of data associated with the smart grid. In this paper, we focus on the issue of privacy linked with consumer information. Potential privacy concerns of smart grid consumers include: how the required information is going to be collected, used and disclosed, how customer information is expected to be safeguarded and how it may be used for or against the consumers; how

permissions will be granted for the collected data to be shared with multiple agencies; and the liabilities related to any breaches of consumer information. It is also worthwhile exploring how the smart grid will "know" about individuals. For example, the energy fluctuation pattern of home appliances is so unique that it may be possible to infer, for example, the model applied for a user's refrigerator. It is also worth noting that many times data that is harmless when collected in isolation may become a privacy threat when combined with other types of data, or examined by a third party for a pattern.

Even when the data about electricity consumption is not collected at regular intervals, information can still be collected at a slower rate through the persistent monitoring of energy consumption. As a result, private information such as how many people live in a household, their presence and absence at home, their schedules for taking showers, watching TV, frequency of microwave use, their sleeping patterns can be collected or deduced. For many individuals, the collection of this type of information represents an invasion of the "sanctity of the home" [5], and one may argue that such intimate details of someone's daily life should not be accessible. The user's data could disclose their usage pattern of electric devices, and very intimate details of household equipment, types, even their possible locations (if the smart grid concept also is combined with smart home concept where, when a person leaves a room, the lights and electric equipment are automatically turned on/off the option of which could be enabled or disabled). In such a case, even the movement pattern of the user within his/her own home could be made!

The privacy concerns discussed above are further confirmed by a recent Privacy Impact Assessment (PIA) conducted by the Privacy Sub-Group of the Cyber Security Working Group [5]. The report has identified the following issues and concerns related to consumer-to-utility information exchanges in the U.S. smart grid:

- There is no clear understanding of the privacy issues on the smart grid.
- There is a lack of standards, privacy policies, or procedures by the entities involved in the smart grid and the collection of information.
- Definitions of personally identifiable information are incomprehensive and inconsistent in the utility industry.
- Smart meters and distributed energy systems may reveal information about residential consumers and activities within the house.
- Roaming smart grid devices (e.g., electrical vehicle recharging at other charging stations such as a friend's house) may generate more personal information.
- Even though the National Association of Regulatory Utility Commissioners adopted the 2000 resolution urging the adoption of privacy principles, only a few State utility level commissions have begun to assess privacy issues associated with the smart grid. This is the case with the State of California through its eight Fair Information Practice (FIP) principles such as transparency, right to access information collected (individual participation), individual access to see and copy information stored on an individual, limited types of information that may be collected on an individual (collection limitation), limited internal use of information about an individual, data quality and integrity, data security, accountability and auditing.

As we mentioned previously, the possibility of learning information about individuals' behaviors, personal habits and lifestyle raises concerns. This becomes an important issue when this information can be used for other purposes besides delivering electricity. Electric utilities and other providers may have access to information about the in-house activities of customers, the times when they are using various devices and appliances as well as the type of devices being used. The initial goal of collecting electricity usage information to generate an elec-

tricity profile has now become a source of behavioral information with an immense potential. The most serious threats related to the privacy deterioration of smart grid consumers include: cyber-attack and intrusion, identity theft, tracking and observing the behavioral patterns of the consumers and the appliances being used, and real time spying and surveillance [20].

3 Proposed Approaches to Protect Privacy in Smart Grid

Given that several schemes have been proposed so far to implement smart grid privacy, our goal in this section is to focus on and compare these recently proposed approaches and architectures aimed at protecting the privacy of smart grid users. Some of these schemes [21] include Anonymous Credential, 3rd Party Escrow Architecture, Load Signature Moderation (LSM), ElecPrivacy, Smart Energy Gateway (SEG), Privacy-preserving Authentication, among others.

In [22], the authors consider a smart grid network as three basic layers: at the highest layer, there is a control center maintained by the power operator, the second layer has substations inside the distribution network and each substation is responsible for the power supply of an area and the lowest layer has the smart meters which are placed at the users' premises as shown in Fig. 4. The proposed Anonymous Credential architecture [22] preserves users' privacy information, including their daily electricity usage pattern from third parties as well as from the power operator. The scheme is based on blind signatures. Blind signature is a method that allows the first party (Party 1) to sign a message generated by a second party (Party 2), without knowing its actual content. When a third party (Party 3) receives the signed message, it can verify that the message is signed by Party 1. The Anonymous Credential scheme uses the blind signature technique to allow the control center (Party 1) to sign a credential generated by a customer (Party 2) without knowing its actual content. At a later time, the control center itself (Party 3) can verify that the credential is indeed signed by Party 1 without knowing who requested the signature or when the signature was generated. The usage of the blind signature technique in this scheme is as follows: The customers prepare a set of credentials, each stating the amount of electricity requested, and request the control center to sign them blindly so that the customer can submit any of these credentials for the request of electricity. Since Party 1 does not know the actual content of the message sent by Party 2, the message is verified using a special technique which is widely adopted in e-cash schemes. Party 2 generates n messages using different blinding factors. It then blinds the n messages and sends them to Party 1. Next, Party 1 randomly chooses m messages ($m < n$) and challenges Party 2 to reveal them by providing the m blinding factors. If the m blinding factors are correct, Party 1 accepts the signature request and signs the remaining $(m - n)$ messages. The scheme assumes that any smart meter can communicate with the control center via a secure communication channel (such as one using the Advanced Encryption Standard (AES) and third parties cannot read the contents without the key concerned).

When a customer presents a credential anonymously, the control center cannot tell which customer is making the request, yet it can verify the signature to confirm that it is from a valid customer (since only valid customers can request blind signatures). The four phases involved in the Anonymous Credential scheme are as follows:

- *Setup phase:* the control center assigns itself a Ron Rivest, Adi Shamir and Leonard Adleman (RSA) public and private key pair for signing credentials.
- *Registration phase:* carried out at the beginning of each month. This phase is *not anonymous*. Customers need to be authenticated using their real identities via an authenticated channel.

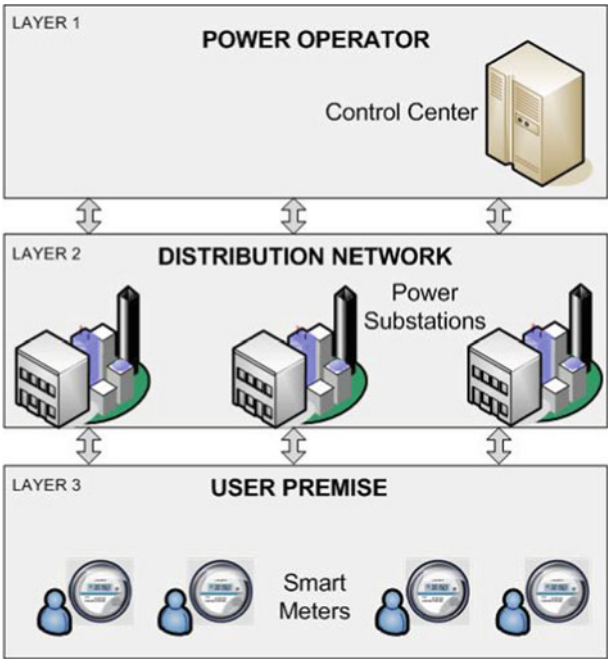

Fig. 4 A 3-layer smart grid system [22]

- *Power requesting phase:* can be executed at any time during the month when the smart meter of a customer finds that it needs more power to support all the electric appliances. This phase is *anonymous*. Customers are validated via anonymous credentials.
- *Reconciliation phase:* carried out at the end of each month. This phase is not *anonymous*. The smart meter sends the unused credentials back to the control center to evaluate the amount of power requested so far.

The 3rd Party Escrow Architecture [20] provides a mechanism for anonymizing high-frequency energy measurement data (such as usage patterns of specific electrical appliances) through the use of a pseudonymous identity (ID). The anonymous meter readings are difficult to associate with a particular smart meter or customer, thus offering a higher level of privacy to the smart grid user.

The distinguishing feature of the Escrow smart meter is that it has two separate IDs, rather than a single ID as is the case with standard smart meters. The two IDs are the High-Frequency ID (HFID) which is anonymous, and the Low Frequency ID (LFID) [23], which is attributable (can be related to a specific customer/smart meter). The main idea of the scheme is to provide anonymity of the HFID messages. The anonymity is implemented by not disclosing the HFID to the utility or the smart meter installer. The HFID is 'hidden' inside the smart meter, or hard-coded to be used for all HFID-related messages. In order for the utility to verify the legitimacy of the HFID, a 3rd party Escrow mechanism is implemented. The 3rd party can be the manufacturer of the smart meter itself or some other trusted 3rd party which has been given access to this information. The manufacturer can assign two unique IDs to each smart meter that is produced, only one of which (LFID) is visible to the utility, both during the procurement and deployment procedures. Essentially, the manufacturer (or the Escrow service) is the only party which is aware (and has a record) of the connection between a valid

HFID/LFID pair. The Escrow is required to comply with a strong data privacy policy. For example, the Escrow may be not expected to access, process or store smart metering data—it will only know about the relationship between a valid HFID and LFID.

The LSM scheme [24] suggests that the home electrical power routing can be used to moderate the home's load signature in order to hide appliance usage information. Load signature is defined as a series of time-stamped average power loads $p(t)$ derived from cumulative energy values $e(t)$ metered at intervals Δt; $p\,(t) = \frac{e(t)-e(t-\Delta t)}{\Delta t}$. A '*home load signature'* is the sum of all home appliance loads. For performing load signature moderation, the authors assume that future smart homes will contain a variety of energy storage and energy generation devices, and thus '*electrical power routing*' will be feasible. Electrical power routing means the selective control and power mixing of a number of electricity sources to '*route*' electricity to a number of consumers. For instance, a kettle drawing 2 kW of power when switched on; the power router could be configured so that 1kW is supplied from a solar panel, 0.5kW from a battery, and 0.5kW from the main electricity supply. The basic contribution of this paper is that it presents the idea how to provide sufficient privacy for the user by including privacy mechanisms for the smart meters which is supposed to record the usage. The authors also propose a power management model using a rechargeable battery, a power mixing algorithm, and evaluate its protection level by proposing three different privacy metrics: an information theoretic (relative entropy), a clustering classification, and a correlation/regression one. We briefly review these metrics below.

Relative entropy: the relative entropy or Kullback Leibler distance [25] is a well-known information theoretic quantity which can be used to compare two sources of information. The distance here is not the mathematical meaning of distance but rather it quantifies the relation between probability densities. If p_0 and p_1 are two probability densities, the Kullback-Leibler distance is defined to be,

$$D(p_0||p_1) = \int_{xmin}^{xmax} p_1(x)\log\frac{\mathrm{p_1(x)}}{\mathrm{p_0(x)}}\mathrm{dx}$$

where, $p_0(x)$ and $p_1(x)$ are the Probability Density Functions (PDFs) of p_0 and p_1, respectively.

> Relative entropy is always positive, and for identical p_0 and p_1, it is zero. Hence, the authors in [26] state that the level of privacy protection offered by a mapping $\emptyset$ can be measured by the relative entropy, $D_\emptyset(p_0||p_1)$ such that the higher the level of protection offered by $\emptyset$, the larger the relative entropy.

Clustering classification: the authors propose using any of the available clustering classification mechanisms which takes a set of data with a distance metric and group them into n clusters that minimize the distance between points. The distance metric here is the difference between power consumption values. They propose to use a simple method of trace analysis that aims to recover information about device power usage from less information sent via the signals.

Regression analysis: as a third metric, the work described in [26] quantifies privacy by combining cross correlation and regression procedures, which can be termed as '*regression analysis*'. In statistics, regression analysis includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables. A dependent variable is what is measured in an

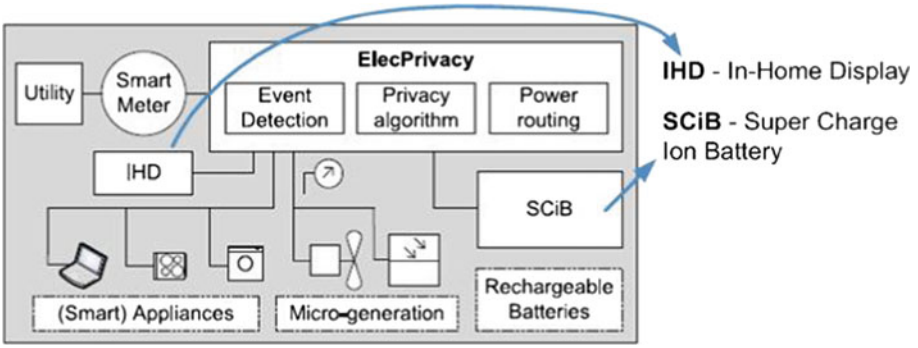

Fig. 5 Components of the ElecPrivacy system as presented in [28]

experiment and what is affected during the experiment. This kind of variable responds to the independent variable. It is termed so because it "*depends*" on the independent variable. In a scientific experiment, there cannot be a dependent variable without an independent variable. Just as an example, if someone is interested to find out how time spent for studying changes "*test score*", then it is understood that the test score does not change time spent for studying, as that had happened earlier. In this case, "studying time" is independent variable and "test score" is dependent variable. Based on these foundations and ideas, the authors in this work apply regression analysis on the received signals to recover information by comparing them over time.

This work can be extended to include other types of privacy metrics such as mutual entropy, or equivocation, introduced in [27]. Also, '*smarter*' battery privacy algorithms may be designed, which the authors have left as future works.

The authors of [28] address some of the issues that were unresolved or unanswered in the work presented in [26]. This work basically is an extension of the authors' previous work in which they study the cost of using rechargeable batteries. Energy usage data directly collected from smart meters could expose lots of information about the user's possessions and types of equipment that he may use. Hence, one way to hide the usage data (i.e., ensuring some kind of privacy) is to use rechargeable battery instead of direct collection of readings from the smart meter. In this case, the cost of the battery becomes a critical factor. A good performing and dynamic rechargeable battery may be costly or replacing it may demand sufficient amount of money. Hence, for such kind of privacy protection, a cost effective solution is needed. This work basically targets this particular issue side-by-side presenting some more thoughts on privacy protection. The authors propose some privacy protection algorithms that can help reduce the exposure of sensitive energy usage information. Their analysis of simulated metering data and real data gathered from an apartment showed that; expected consumption events can be predicted quite well, rechargeable battery resources may protect privacy of particular sets of appliances, carefully chosen batteries can last a long life, and the impact on the utility is positive. They also show that the initial system operational costs suggest that their ElecPrivacy system (illustrated in Fig. 5) offers other benefits besides privacy protection, such as improved load balancing.

ElecPrivacy system has four sub-systems included in it:

- *Metering mechanism*: this is used to obtain a set of electricity measurements from the smart meter or from smart appliances.

- *Event detection*: this subsystem analyzes metering data in order to detect an occurring, or predict an imminent, event that may contain '*privacy information*'. For example, this may be a power trigger generated by a particular event, such as a change in power consumption (e.g., appliance switch-on/off event).
- *Privacy protection algorithm*: it configures power routing to mask a detected consumption event. Different protection settings may be edited with the help of an in-home display (IHD).
- *Power routing*: it mixes a private (i.e., non-utility) energy resource (e.g., rechargeable battery) with utility energy to meet appliance demands.

The authors note that the ElecPrivacy system may also be implemented within a '*charge grid system*' [29] that uses a Super Charge Ion Battery (SCiB^{TM}) pack [30] and a bidirectional inverter to optimize the flow and storage of electricity. Optionally, ElecPrivacy may also control energy generated locally from photovoltaic (PV) panels or wind turbines (micro-generation).

ElecPrivacy system is expected to detect a privacy threat and accordingly respond by configuring power routing to hide appliance load signatures. The same definitions of power routing and load signature as presented earlier are meant by these terms here. Hence, in this system, privacy threats may be detected either at the time of their appearance or in advance (e.g., by following the analysis of scheduled, desired, or predicted future events within the home) and accordingly selective control and power mixing of a number of electricity sources can be performed to cover consumption demands.

The SEG architecture [31] is deployed at the user premises and uses a privacy manager, which is designed as a software component running on SEG, deployed at users' premises. The idea of the work is to provide user centric privacy that is, the user could be in control of own privacy parameters. The proposed privacy manager has the ability to specify privacy conditions and obligations with respect to the handling of users' private data, and to rely on SEG security architecture features such as application isolation, mandatory access control, pseudonymity, and secure storage to reliably enforce the users' specified privacy constraints. The main features of the privacy manager are as follows:

- *Customer privacy preferences specification and enforcement*: the energy customer would express how revealed personal information should be handled and the utility or service provider would express how costumer's information will be treated.
- *Privacy policies enforcement*: each SEG application policy is bound to a smart software agent and has to be validated against the SEG platform integrity policy both during the installation and at runtime. This ensures that SEG only hosts and runs smart software agents which meet predefined gateway security requirements; e.g., that the former (will not) access locally stored energy usage data collected at this particular premise.
- *Secure storage and data masking*: the secure storage will guarantee the confidentiality and accuracy of locally stored energy usage data. Only trusted and legitimate applications (e.g., billing provider software agent) can access the metered data repository.
- *Pseudonymity*: enables the customer to use smart grid resources or related services without revealing their respective identities but remaining accountable for their transactions.
- *Privacy feedback*: allows the display of feedbacks to the energy customer, regarding the handling of its personally identifiable information.

The Privacy-preserving Authentication Scheme for a smart grid network (PASS) [32] involves the use of smart appliance (located at customers' homes) attached with a tamper-resistant device for generating pseudo identities and signatures on messages. A customer is given this

device when he/she opens an account or registers a newly purchased smart appliance. The characteristic features of the PASS architecture are as follows:

- *Message authentication:* before a smart appliance transmits a request message to the control center, it has to include a Hash-based Message Authentication Code (HMAC) signature on the message using the regional system key. This regional system key is only known by the control center, the substation and all tamper-resistant devices within the region. Hence, an outside attacker (who does not belong to the region or is not a registered smart appliance) does not know how to generate a valid HMAC signature. Thus, the PASS scheme protects from outsider attacks.
- *Identity privacy:* in all request messages sent by a smart appliance, pseudo identities instead of real identities are used.
- *Request message confidentiality:* the amount of electricity required by a smart appliance is encrypted using the public key of the control center. Thus, except for the control center, no one can decrypt the value representing the electricity amount. On the other hand, the encryption feature in the PASS architecture allows a substation to aggregate request messages sent by smart appliances within its region but the substation does not need to know about those individual amount values.

The work in [33] analyzes security and privacy in smart grid and specifically emphasizes the privacy aspects. The authors propose a secure and efficient in-network data aggregation and dispatch scheme for AMI in home area networks for the smart grid. In-network aggregation is the process of collecting content from multiple sources or devices in a network. With this mechanism, the authors propose adopting *Walsh function* based on Hadamard code to generate mutual orthogonal chip codes to be used in the secure in-network data aggregation and dispatch scheme. The use of orthogonal code allows multiple users to communicate simultaneously over a single frequency. This is achieved by the use of spreading codes, whereby a single data bit is "*spread*" over a longer sequence of transmitted bits. These codes, also known as chip sequences, must be carefully chosen so that the data may be correctly "*dispread*" at the receiver. Such codes are known as orthogonal codes. The Hadamard code [34] is an error-correcting code that is usually used for error detection and correction when transmitting messages over very noisy or unreliable channels. In their work, the authors apply these techniques envisioning that the smart meter works as an authentication server that is connected with multiple smart devices and each smart device contributes to the formation of confidential data which can be regenerated at the smart meter. This work describes the coding techniques and the steps on how the original data readings are spread and then mixed up with the spreading code of other smart devices. The smart meter can reconstruct the original reading data from the mixed data using the chip code established with smart devices in their initialization procedure through mutual authentications.

The authors of [35] present a slightly different perspective of smart grid than the traditional view which is a bi-directional supply chain linking power generation to transmission, distribution, and consumers using information communication technologies. The authors consider it not as a separate system but as a Cyber-Physical System (CPS) that blurs the line between physical electricity infrastructure and Cyber-infrastructure, with the Internet providing the backbone for utilities to control operations and communicate with consumer appliances. A CPS can be considered as a combination of networked embedded systems and physical environments. Recent CPS research efforts have focused on addressing integration issues resulting from networked embedded systems together with their surrounding environment. In the field of CPS, multiple embedded control systems interact among themselves through communications and physical environments (e.g., from actuators to sensors). Thus, the integration

of multiple embedded controllers with physical systems may become a very challenging task. Hence, considering smart grid as a CPS and putting its privacy issues under '*CPS privacy*' becomes another important issue to consider. However, the work presents a good analysis of the system's security and privacy where various user characteristics, data characteristics, application characteristics, and platform characteristics are discussed in detail.

As experimentations were being done alongside deployment of smart meters in users' premises, the issue of end-user privacy becomes a critical research topic. While end-user privacy is a real concern to get a greater acceptance of smart grid technologies among users, there is another kind of competitive privacy problem that arises at the level of Regional Transmission Organizations (RTOs) because of the conflicting objectives of sharing data for distributed estimation and blocking data for economic (competitive) and end-user privacy reasons. Hence, there should be some kind of trade-off between how much data could be shared and how much can be withheld or controlled to ensure profitability and privacy. The work in [36] focuses on this issue and proposes an information-theoretic approach for competitive privacy in smart grid. The work is based on mathematical modeling and formulations of concepts. The authors present a mathematical model for the grid at the level of the RTOs that takes into account the interconnections amongst them. Viewing the power system state at each RTO as an information source, they model the measurements at each RTO as a linear combination of all the sources.

The work in [37] provides a statistical method based on Empirical Probability Distribution (EPD) to analyze the content of power signals from the viewpoint of privacy. The authors propose a technique based on the EPD which is applied for studying two kinds of signals; one whose privacy is not protected and second type that uses privacy algorithms to hide data. The difference between classical and empirical probability is that classical probability assumes that certain outcomes are equally likely, while empirical probability relies on actual experience to determine the likelihood of outcomes. Empirical probability usually estimates probabilities from experience and observation [38].

The authors of [39] discuss the interactions among the actors of the future smart grid infrastructure. An actor is defined as an entity which has a role in a system and is one of the parties involved in the smart grid infrastructure. The privacy issue contribution of the work described in [39] suggests that, in addition to the use of '*anonymization*' of the measurement data sent to the smart grid Distribution Service Operators (DSOs), it is important to consider the whole data treatment chain as sensitive. It means that at each step where an actor is involved in dealing with the data (from user to the DSO or among all entities), that should be considered confidential and treated as highly restrictive.

The authors of [40] propose a cooperative state vector estimation technique that preserves the privacy of the personal behavior of the user. The key objectives are to ensure mainly two things: (a) the power consumption measurement is well obfuscated such that users do not fully disclose their private behavioral information, and (b) the obfuscated data retain the necessary or basic information such that the state vector (a column vector whose components are the state variables of the system) can be accurately estimated from the perturbed data. "*Perturbed data*" is the original measurement data that is perturbed to conceal it and to make it difficult to infer the original data. Another significant contribution of this work is that the authors evaluated the performance of the proposed data obfuscation scheme with 1349 measurement data sets. For this, they used the data sets as if they are connected to 5 different IEEE Test Systems that are portions of the Middlewestern U.S. Electric Power Grids. They also evaluated the illegibility to human inspectors, resilience to automated data mining attackers, and communication overhead.

The work in [41] presents a holistic privacy engineering approach for smart grid systems. The authors analyze various privacy issues in future energy systems, discuss privacy-aware design methodology and countermeasures to protect privacy.

The authors of [42] propose an Energy Privacy Preserving Aggregation (EPPA) scheme for secure smart grid communications. It presents a multi-dimensional data aggregation approach based on the homomorphic Paillier cryptosystem [43] which is composed of three algorithms namely, key generation, encryption, and decryption. The proposed technique is based on composite residuosity classes, whose computation is believed to be computationally difficult. It is a probabilistic asymmetric algorithm for public key cryptography and inherits additive homomorphic properties. Homomorphic encryption allows specific types of computations to be carried out on ciphertext and obtain an encrypted result. For example, one user could add two encrypted numbers and then another user could decrypt the result, without either of them being able to find the value of the individual numbers. Homomorphic encryption schemes are malleable by design [44]. A more in-depth discussion of Paillier cryptosystems can be found in [45,46]. Many of the existing data aggregation schemes [47–49] collect information as one-dimensional information. However, smart meter data could be considered as multi-dimensional in nature, because, these include including various aspects of the information such as the amount of energy consumed, the time it was consumed, the purpose of the consumption, and so on. Considering the high data collection frequency, multi-dimensional information and the large number of users, current data aggregation schemes generate not only huge communication costs but also impose overwhelming process load on local gateways. In contrast to traditional one-dimensional data aggregation methods, EPPA is shown to significantly reduce computational cost and significantly improve communication efficiency, satisfying the real-time high-frequency data collection requirements in smart grid communications. The main drawback of the work is that it is highly theoretical and it does not provide enough details on how such an approach can be deployed in practice.

The authors of [50] outline Grid 2.0 Research, a collaborative smart grid research program between Gachon Energy Research Institute (GERI) of Kyungwon University, South Korea and Bell Labs of Alcatel-Lucent. They discuss economic modeling, networking, security, and privacy issues in smart grid. The contribution of [50] towards privacy issues is limited to the fact that the authors consider a smart grid scenario as a client-server model, where the clients have information to be communicated with the server and the clients' information need to be kept confidential. The authors consider mainly two points related to privacy: (a) detailed electricity consumption measurements, which are invasive enough to allow identification of appliance brands in the household, and (b) the electric vehicles charging, which reveals person's location and distances travelled. The authors envision that plug-in batteries would be used for Electric Vehicles (EVs) as power sources when EVs are stationary but connected to the power grid. Since they consider the smart grid infrastructure as a client-server model, they propose the use of a third party data proxy for ensuring privacy. This is a common solution [51–53] used in systems such as credit card authentication or any other client-server network setting where the proxy agents collects the data, anonymizes it, and sends it for batch processing to the server. The proxy agent can also participate in the server-to-client communication without colluding with the server.

The authors of [54] deal with preserving the privacy of metered data. The authors propose a set of privacy-preserving protocols amongst a provider, a user agent and a simple tamper-evident meter. This work considers a scenario illustrated in Fig. 6. As shown in the figure, the privacy of the metered data is preserved by employing encryption mechanisms along with certification techniques. Within the boundary shown in the diagram (i.e., home environment)

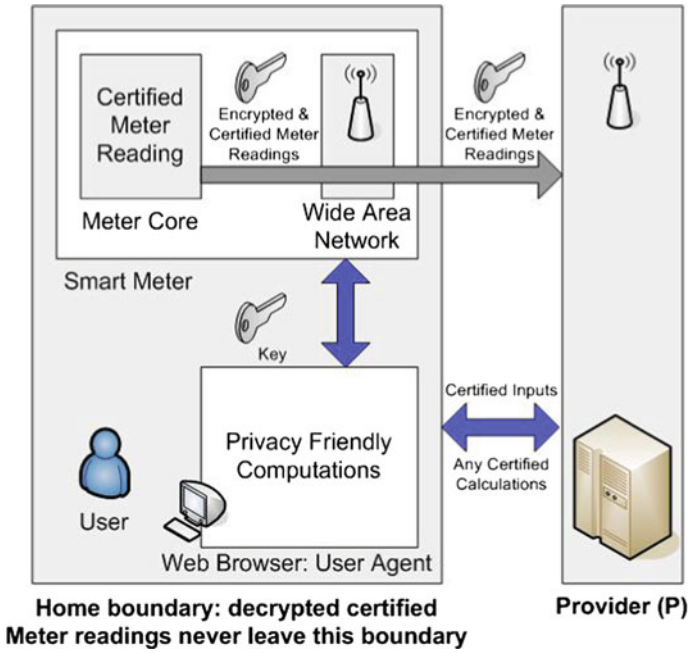

Fig. 6 Interaction among various parties involved in a smart metering scenario [54]

plaintext is used but when sending or communicating with entities outside the boundary, certification and encryption techniques are used. The authors argue that their scheme can be applied to all types of smart metering including electricity, water and gas metering, and can be extended for other future smart meter based systems. The main contribution of this work can be summarized as: the meter produces certified readings of measurements and transmit them to the user via a secure communication channel. For billing, the user combines those readings with a certified tariff policy, to produce a final bill. The bill is then transmitted to the provider alongside a *zero-knowledge proof* that ensures the calculation to be correct and leaks no additional information. A *zero-knowledge proof* of knowledge [55] is a two-party protocol between a prover and a verifier. The prover demonstrates to the verifier its knowledge of some secret input (witness) that fulfills some statements without disclosing this input to the verifier. The protocol should meet two properties: (a) it should be a proof of knowledge; that means, a prover without knowledge of the secret input convinces the verifier with negligible probability, and (b) it should be zero-knowledge; that is, the verifier learns nothing but the truth of the statement. The fact that a witness is not distinguishable is a weaker property which requires that the proof does not reveal the witness (among all possible witnesses) used by the prover. This work gives the user some choice in terms of how to deal with the data readings.

The authors of [57] develop a theoretical framework that abstracts both the privacy and the utility requirements of smart meter data. The authors assume that actual (i.e., real) load measurements are sampled (at an appropriate frequency) from a smart meter, and can be correlated (models the temporal memory of both appliances and human usage patterns). As the authors claimed in this work, the theoretical framework allows for the precise quantification of the utility-privacy tradeoff problem in smart meter data. This work provides a strong

theoretical foundation (in terms of theorems, corollary, definitions, and theoretical proofs) that addresses the utility-privacy tradeoff problem which is about privacy preservation and the need for precise electricity usage information in order to enable the smart distribution of electricity. The real impact of this proposed model remains to be seen through an actual implementation. Table 2 summarizes the benefits, limitations, and ease of implementation of the privacy architectures and approaches discussed above.

It is worth pointing out that many of the smart grid privacy architectures discussed in this section incur different computational costs in their operations. A major factor that will determine the actual success of these architectures will, to a large extent, depend on their ease of implementation (hardware or software) in practice as highlighted in Table 2.

3.1 Discussion

Privacy concern is an important issue in many areas such as electronic voting, wireless sensor networking, online banking applications where the user needs to be authenticated. Various privacy protection schemes have been proposed to prevent customer identification; they vary depending on the context and the architecture in use. In the case of electronic voting, several schemes had been developed in order to maintain the voter's privacy and to make it impossible for anyone to determine how each voter voted. Such schemes are mainly based on the use of either blind signatures [62] or homomorphic encryption-based [63]. Most of the solutions discussed earlier had been proposed to prevent the identification of customers based on the application-layer information. However, they do not take into consideration the fact that lower layers (such as the link and network layers) of the communication stack may reveal identity information. For example, reusing the same IP address over time may eventually allow an attacker to perform traffic analysis to identify the customers. It is worth noting that different solutions (such as Crowds [64]) had been developed to prevent user identification at the network and link layers. The approach used in Crowds operates by grouping users into a large and geographically diverse group (crowd) that collectively issues requests on behalf of its members. Hence, the servers are unable to learn the true source of a request because it is equally likely to have originated from any member of the crowd, and even collaborating crowd members cannot distinguish the originator of a request from a member who is merely forwarding the request on behalf of another [64]. Since communications between smart meters and energy suppliers is based on IP, Crowds could be used as a solution to protect the information related to the smart meters.

Privacy in wireless sensor networks may be broadly classified into two categories [65]: content privacy and contextual privacy. The issue of contextual privacy arises due to the nature of wireless communication media that can expose contextual information about the encrypted content being transmitted by the network. Different approaches have also been designed to protect the user privacy in location tracking systems and are discussed in [66]. It is worth pointing out that customer location privacy is not normally considered as a serious issue by the smart grid privacy solutions reviewed earlier.

In the case of client/server architectures (such as those used by online banking transactions), the server needs to identify the client before allowing access to the desired service. The client sends its identity information (e.g., login credentials or certificate) to the server and relies on lower-layer security protocols (TLS (Transport Layer Security), SSH (Secure SHell), IPSec (IP Security)) to send the identity information securely to the destination without being disclosed to unwanted parties. In the case of smart grid, this is not a complete solution because the energy supplier will be able to identify the customer and its requested content as well.

Table 2 Benefits, limitations, and ease of implementation of current smart grid privacy architectures

Privacy architecture	Ref.	Privacy protection in smart grid		
		Benefits	Limitations	Ease of implementation
3rd Party escrow	[20]	Privacy via anonymyzing frequently reported information	Provision of attributable metering data is not precluded	Low
Credential-based privacy-preserving power request	[22]	Efficient, low communication overhead	Credential identity collision (two or more credential registrations with the same credential identity) cannot be avoided	Low
Load Signature Moderation (LSM)	[24]	Independence to the operational functionalities of the smart grid	Unsuitable operational performance in terms of offering privacy	Medium
Smart Energy Gateway (SEG)	[31]	Intelligent computer software acting on behalf of their users can optimize interaction with the utility or in-premise smart appliances	Only a preliminary model and theoretical framework that are yet to be validated by an actual implementation of the protocols and software of the SEG architecture	Medium
Privacy-preserving authentication	[32]	Privacy of customers (including their daily electricity usage) can be preserved while at the same time the control center can generate and distribute a proper amount of electricity	Increased delays for authentication of packets at substations	Low
Secure data aggregation and dispatch scheme	[33]	Only the smart meter can reconstruct the original data reading from the mixed data using the chip code established with smart devices in their initialization procedure through mutual authentications	A preliminary work in which optimization issues are is not analyzed for home power management systems with regards to the privacy of customer power usage behaviors	Medium

Table 2 Continued

Privacy architecture	Ref.	Privacy protection in smart grid		
		Benefits	Limitations	Ease of implementation
Extension of LSM (ElecPrivacy)	[28]	Cost-effective battery solution for privacy for home smart meters. It is possible to hide appliance events and save money for peak load shedding contributions	Detailed cost analysis is not done properly, and which may vary depending on the application scenario and company/user requirements	Low
Analysis of security and privacy issues in smart grid on clouds	[35]	Different perspective of smart grid than the traditional view which is a bi-directional supply chain linking power generation to transmission, distribution, and consumers using information communication technologies	Integration of multiple embedded controllers with physical systems may become a very challenging task	Low
Information theoretic approach for smart grid privacy	[36]	A theoretical model for formalizing the conflicting objectives of estimation accuracy and competitive privacy in smart grid operations	Theoretical model is yet to be validated by an actual implementation or the practical use of the model is not verified	Low
Data mining and privacy of personal behavior	[37]	Use of EPD to analyze the content of power signals from the viewpoint of privacy	Privacy of different kinds of EPD databases are yet to be analyzed using the notion of differential privacy	Low
Security and privacy in smart grid architectures	[39]	Discusses the interactions among the actors of the future smart grid infrastructure	The power supply infrastructure and smart grid actors considered by the authors may not be universally recognized	Low
Preserving privacy of user behavior in smart grid	[40]	Propose a cooperative state vector estimation technique that preserves the privacy of the personal behavior of smart grid users	Although practical usage data are used (Middlewestern U.S. Electric Power Grids) for providing empirical evidence, more rigorous privacy guarantees remain to be studied	Medium

Table 2 Continued

Privacy architecture cline3-5	Ref.	Privacy protection in smart grid Benefits	Limitations	Ease of implementation
Efficient and privacy-preserving aggregation scheme	[42]	Use of the EPPA scheme for secure smart grid communications	EPPA is not very specific only for smart grid but rather could be applied to other networks such as WSNs or wireless mesh networks, for which no proper justification is provided to limit it only to smart grid scenario	Medium/ high
GERI	[50]	Collection of anonymous data and its send for batch processing to the server through a proxy agent	Basic idea is not very specific only for smart grid. Similar methods are used often for other systems such as credit card authentication or any other client-server network setting	Medium
Privacy-preserving smart metering	[54]	The proposed scheme can be applied to all types of smart metering including electricity, water and gas metering	Proposed mechanisms rely on zero-knowledge proof between a prover and a verifier	Medium
Utility privacy framework	[57]	Presents a theoretical framework that abstracts both the privacy and the utility requirements of smart meter data	The real impact of this proposed model remains to be seen through an actual implementation	Medium

Content privacy could also be ensured through a privacy proxy [67] (also known as a trusted third party). However, there would be a privacy issue if the proxy acts maliciously and shares with the energy supplier the (signed) customers' requests. An advanced solution based on the use of an additively homomorphic encryption is presented in [68]. Instead of using a proxy, the solution proposes forming multiple groups of smart meters; each group is formed of several smart meters belonging to the same building/street and is limited to one energy supplier. One smart meter is periodically designed as the key aggregator that will collect the other customers' requests—encrypted within an additively homomorphic encryption—and then sends them to the energy supplier. However, if a smart meter sends its key without its encrypted content (or vice and versa), the energy supplier will not be able to decrypt the aggregate value received from that smart meter group and the whole process will fail. Hence, the authors propose adding a token-based solution which increases the complexity of the proposed smart grid privacy approach and makes it difficult to validate in a real deployment.

NIST has already articulated the privacy challenges and recommendations for the smart grid [60]. However, we found that many of the recently proposed privacy solutions and architectures described in the literature do not always follow these recommendations. While reviewing these aforementioned solutions we also found that some efforts have attempted to adapt privacy solutions that have been in use in other application domains (such as wireless sensor networking, etc.) to the smart grid ecosystem. Such approaches face serious design and implementation challenges because of the inherent characteristics of the smart grid technology and its environment. The authors of [69] have proposed an interesting methodology to help identify and deal with privacy and data protection challenges throughout the engineering phase of the smart grid. To address the privacy challenges in smart grid, we need to adopt a holistic approach that can provide a privacy protection solution that is simple, scalable, cost-effective, and incurs minimal computational processing/communication overheads.

4 Government Regulations for Smart Grid Privacy

The first officially reported instance of a privacy intrusion into the smart grid in the U.S. occurred in April 2009. It was found that spies made random and successful attempts to infiltrate the smart grid in order to cause much more severe disruption in the future [58]. This incident led the authorities to develop and implement the legislature and regulations needed to address privacy issues of the smart grid while maintaining the reliability and efficiency of the technology. In the U.S., smart grid technologies are likely to require a dual policy of privacy and disclosure [59]. A cyber-attack on a power grid could involve unauthorized access, alteration, deletion and/or theft of data. In this case, law enforcement agencies in the U.S. will be expected to investigate these attacks and other crimes using utility evidence. The smart grid privacy issues require detailed studies and public input so that current laws can be adapted and new laws can be created. Presently, there are no well-defined laws and regulations to protect privacy on the smart grid. It is assumed that current privacy laws such as the Gramm-Leach-Bliley Act, the Children's Internet Protect Act (CIPA), the Electronic Communications Privacy Act (ECPA), and protection provided by privacy in the home laws (fourth and fourteenth amendments of the U.S. constitution) can be modified to deal with privacy issues associated with smart grid technology use. There are three legal approaches to protect privacy [60] in smart grid:

- *Constitutional protection*: covering personal communication and activities.

- *Data-specific protection*: covering specific items such as credit card numbers and social security numbers, or specific technology such as computers used for data storage.
- *Contractual protection*: outlined for business contracts.

It is worth mentioning that the focus of smart grid technology in Europe is primarily as an exclusive source of renewable energy production (in contrast to the U.S. view which focuses on the convenience of the local consumers). The Italian ENEL Telegestore Project is accredited as the first and the largest and smart grid Project of Europe. The smart grid Industry within Europe lacks the availability of legislatures and the regulations for smart grid privacy, unlike the United States where the regulatory agencies such as the National Association of Regulatory Utility Commissioners' (NARUC) and other government agencies are collaborating to ensure the privacy of the consumers of smart grid [61].

5 Conclusion

In the last couple of years, we have witnessed huge investments and interests from industry and governments in smart grid technologies. Various stakeholders (residential/commercial customers, local government, utility operators, etc.) are expected to reap several benefits associated with the smart grid which include: improved energy efficiency, increased reliability, reduced energy costs, greater flexibility in energy consumption, better safety and security, and an improved environment (through renewable, renewable non-variable, non-renewable/non-variable energy sources). The deployment of smart grid technologies has also generated considerable interests in data privacy issues of smart grid users. The privacy concerns are mostly related to the collection and use of energy consumption data. In this context, we have discussed various smart grid privacy issues and we have presented various smart grid privacy architectures and approaches that have been recently proposed in the literature. We also identified the various strengths and weaknesses of these privacy solutions. The success of smart grid technology and its wide acceptance rely on gaining the trust and confidence of customers which in turn depend on assurances regarding the protection of their privacy.

Acknowledgements The authors would like to thank the anonymous reviewers for their suggestions and feedback which helped us to improve the quality and presentation of this paper. Sherali Zeadally was partially supported by a District of Columbia NASA Space Grant and an NSF TIP grant (Award Number 1036293) during the course of this work.

References

1. McClanahan, R. (2003). SCADA and IP: Is network convergence really here?. *IEEE Industry Applications Magazine, 9*(2), 29–36.
2. Fan, J., & Borlase, S. (2009). The evolution of distribution. *IEEE Power and Energy Magazine, 7*(2), 63–68.
3. Farhangi, H. (2010). The path of the smart grid. *IEEE Power and Energy Magazine, 8*(1), 18–28.
4. NIST. (2012). *NIST framework and roadmap for smart grid interoperability standards, release 2.0*. NIST Special Publication 1108R2, February 2012.
5. Mark, J. (2010). *New electricity grids may be smart, but not so private—The Denver post*. 18 May 2010. Available at: http://www.denverpost.com/business/ci_15106430 (last accessed 2 Oct 2012).
6. U.S. Department of Energy. (2008). *Advanced metering infrastructure, white paper*. NETL Modern Grid Strategy Powering our 21st-Century Economy, February 2008. Available at: http://www.netl.doe.gov/smartgrid/referenceshelf/whitepapers/AMI%20White%20paper%20final%20021108%20%282%29%20APPROVED_2008_02_12.pdf (last accessed 2 Oct 2012).

7. Federal Energy Regulatory Commission. (2008). *Assessment of demand response & advanced metering*. Staff Report, December 2008. Available at: http://www.ferc.gov/legal/staff-reports/demand-response.pdf (last accessed 2 Oct 2012).
8. Massoud, A., & Wollenberg, B. (2005). Toward a smart grid: Power delivery for the 21st century. *IEEE Power and Energy Magazine, 3*(5), 34–41.
9. Lu, Z., Lu, X., Wang, W., & Wang, C. (2010). Review and evaluation of security threats on the communication networks in the smart grid. In *Proceedings of IEEE military communications conference* (pp. 1830–1835).
10. Steven, J., Peterson, G., & Frincke, D. (2010). Smart-grid security issues. *IEEE Security and Privacy*, 81–85.
11. McDaniel, P., & McLaughlin, S. (2009). Security and privacy challenges in the smart grid. *IEEE Security and Privacy, 7*(3), 75–77.
12. Westin, A. (1967). *Privacy and freedom* (p. 7). New York: Atheneum.
13. Miller, J. (2008). Who are you? The trade-off between information utility and privacy. *IEEE Internet Computing, 12*(4), 93–96.
14. Miller, J. (2008). Who are you, part II: More on the trade-off between information utility and privacy. *IEEE Internet Computing, 12*(6), 91–93.
15. Pedersen, D. (1982). Personality correlates of privacy. *Journal of Psychology, 112*, 11–14.
16. Brierley, N. (1992). *The meaning and use of privacy: A study of young adults*. Ph.D. dissertation, The University of Arizona, USA.
17. International Energy Agency. (2011). Technology roadmap: Smart grids. International Energy Agency, April 2011. Available at: http://www.iea.org/papers/2011/smartgrids_roadmap.pdf (last accessed 2 Oct 2012).
18. Kindy, D., & Pathan, A. (2011). A survey on SQL injection: Vulnerabilities, attacks, and prevention techniques. In *Proceedings of 15th IEEE symposium on consumer electronics (IEEE ISCE 2011)*, Singapore.
19. Singh, M. (2002). Privacy for telecom services. *IEEE Internet Computing, 6*(1), 4–5.
20. Efthymiou, C., & Kalogridis, G. (2010). Smart grid privacy via anonymization of smart metering data. In 1st *IEEE international conference on smart grid communications* (pp. 238–243).
21. Siddiqui, F., Zeadally, S., Alcaraz, C., & Galvao, S. (2012). Smart grid privacy: Issues and solutions. In *Proceedings of second international workshop on privacy, security, an trust in mobile and wireless systems (MobiPST 2012)*, Munich, Germany.
22. Cheung, J., Chim, T., Yiu, S., & Li, V. (2011). Credential-based privacy—preserving power request scheme for smart grid network. In *IEEE Global telecommunications conference* (pp. 1–5).
23. Das, S., Kant, K., & Zhang, N. (2012). *Security and privacy in the smart grid. Handbook on Security Cyber-Physical Critical Infrastructure*, Chap. 25. Morgan Kaufmann, February 2012.
24. Kalogridis, G., Efthymiou, C., Denic, S., Lewis, T., & Capeda, R. (2010). Privacy for smart meters: Towards undetectable appliance load signatures. In *First IEEE international conference on smart grid communications*, (pp. 232–237).
25. Johnson, D., & Sinanovic, S. (2012). *Symmetrizing the Kullback-Leibler distance*. Available at http://www.ece.rice.edu/~dhj/resistor.pdf (last accessed 2 Oct 2012).
26. Kalogridis, G., Efthymiou, C., Denic, S., Lewis, T., & Cepeda, R. (2010). Privacy for smart meters: Towards undetectable appliance load signatures. In *Proceedings of the first IEEE international conference on smart grid communications (SmartGridComm)* (pp. 232–237).
27. Shannon, C. (1949). Communication theory of secrecy systems. *Bell System Technical Journal, 28*(4), 656–715.
28. Kalogridis, G., Fan, Z., & Basutkar, S. (2011). Affordable privacy for home smart meters. In *Proceedings of the Ninth IEEE international symposium on parallel and distributed processing with applications workshops (ISPAW)* (pp. 77–84).
29. Toshiba Inc. (2012). *Solar power generation*. Available at http://www.toshiba.co.jp/env/en/energy/solar.htm (last accessed 2 Oct 2012).
30. Toshiba Inc. (2012). Super-charge ion battery (SCiB^{TM}). Available at: http://www.toshiba.com/ind/product_display.jsp?id1=821 (last accessed 2 Oct 2012).
31. Fhom, H., Kuntze, N., Rudolph, C., Cupelli, M., Liu, J., & Monti, A. (2010). A user-centric privacy manager for future energy systems. In *Proceedings of the international conference on power systems technology* (pp. 1–7).
32. Chim, T., Yiu, S., & Li, V. (2010). PASS: Privacy-preserving authentication scheme for smart grid network. In *Proceedings of the international conference on power systems technology*.

33. Yan, Y., Qian, Y., & Sharif, H. (2011). A secure data aggregation and dispatch scheme for home area networks in smart grid. In *Proceedings of IEEE global telecommunications conference (GLOBECOM 2011)* (pp. 1–6).
34. Phelps, K., Rifa, J., & Villanueva, M. (2005). Rank and kernel of binary Hadamard codes. *IEEE Transactions on Information Theory, 51*(11), 3931–3937.
35. Simmhan, Y., Kumbhare, A., Cao, B., & Prasanna, V. (2011). An analysis of security and privacy issues in smart grid software architectures on clouds. In *Proceedings of IEEE international conference on cloud computing (CLOUD)* (pp. 582–589).
36. Sankar, L., Kar, S., Tandon, R., & Vincent Poor, H. (2011). Competitive Privacy in the smart grid: An information-theoretic approach. In *Proceedings of IEEE international conference on smart grid communications (SmartGridComm)* (pp. 220–225).
37. Kalogridis, G., & Denic, S. Z. (2011). Data mining and privacy of personal behavior types in smart grid. In *Proceedings of* 11th *international conference on data mining workshops (ICDMW)* (pp. 636–642).
38. Bluman, A. (1997). Elementary statistics: A step by step approach. McGraw-Hill College, ISBN: 978-0256234305.
39. Barenghi, A., & Pelosi, G. (2011). Security and privacy in smart grid infrastructures. In *Proceedings of 22nd international workshop on database and expert systems applications (DEXA)* (pp. 102–108).
40. Kim, Y., Ngai, E., & Srivastava, M. (2011). Cooperative state estimation for preserving privacy of user behaviors in smart grid. In *Proceedings of IEEE international conference on smart grid communications (SmartGridComm)* (pp. 178–183).
41. Fhom, H., & Bayarou, K. (2011). Towards a holistic privacy engineering approach for smart grid systems. In *Proceedings of 10th IEEE international conference on trust, security and privacy in computing and communications (TrustCom)* (pp. 234–241).
42. Lu, R., Liang, X., Li, X., Lin, X., & Shen, X. (2012). EPPA: An efficient and privacy-preserving aggregation scheme for secure smart grid communications. *IEEE Transactions on Parallel and Distributed Systems, 23*(9), 1621–1631.
43. Paillier, P. (1999). Public-key cryptosystems based on composite degree residuosity classes. In *Proceedings of Eurocrypyt* (Lecture Notes in Computer Science), Vol. 1592 (pp. 223–238). Berlin: Springer.
44. Rivest, R. (2012). *Lecture notes 15: Voting, homomorphic encryption*. Available at: http://web.mit.edu/6.857/OldStuff/Fall02/handouts/L15-voting.pdf (last accessed 2 Oct 2012).
45. Menezes, A., Oorschot, P., & Vanstone, S. (1996). *Handbook of applied cryptography*. CRC Press, ISBN: 0-8493-8523-7.
46. Stavroulakis, P., & Stamp, M. (2010). *Handbook of information and communication security*. Springer, ISBN 978-3-642-04116-7.
47. Castelluccia, C., Chan, A., Mykletun, E., & Tsudik, G. (2009). Efficient and provably secure aggregation of encrypted data in wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN), 5(3).*
48. Westhoff, D., Girao, J., & Acharya, M. (2006). Concealed data aggregation for reverse multicast traffic in sensor networks: Encryption, key distribution, and routing adaptation. *IEEE Transactions on Mobile Computing, 5*(10), 1417–1431.
49. Shi, J., Zhang, R., Liu, Y., & Zhang, Y. (2010). PriSense: Privacy-preserving data aggregation in people-centric urban sensing systems. In *Proceedings of IEEE INFOCOM* (pp. 1–9).
50. Budka, K., Deshpande, J., Hobby, J., Kim, Y., Kolesnikov, V., & Lee, W., et al. (2010). GERI—Bell labs smart grid research focus: Economic modeling, networking, and security & privacy. In *Proceedings of first IEEE international conference on smart grid communications (SmartGridComm)* (pp. 208–213).
51. Veeravalli, B. (2003). Performance analysis of a generic proxy-based client-server system for World-Wide Web services using a generalized Markov chain model. *Journal of High Speed Networks, 12*(3–4), 111–131.
52. Yang, T., Xiong, H., Hu, J., Wang, Y., Xin, W., Deng, Y. et al. (2011). A traceable privacy-preserving authentication protocol for VANETs based on proxy re-signature. In *Proceedings of 8th international conference on fuzzy systems and knowledge discovery (FSKD)*, Vol. 4. (pp. 2217–2221).
53. Kadowaki, K., & Fujita, S. (2009). A dynamic user management in networked consumer electronics via authentication proxies. In *Proceedings of international conference on parallel and distributed computing, Applications and Technologies* (pp. 195–200).
54. Rial, A., & Danezis, G. (2011). Privacy-preserving smart metering. In *Proceedings of the 10th annual ACM workshop on privacy in the electronic society (WPES '11)* (pp. 49–60).

55. Bellare, M., & Goldreich, O. (1993). On defining proofs of knowledge. In *Proceedings of CRYPTO '92* (Lecture Notes in Computer Science), Vol. 740 (pp. 390–420). Berlin: Springer.
56. Perrig, A., Szewczyk, R., Wen, V., Culler, D., & Tygar, J. (2002). SPINS: Security protocols for sensor networks. *Wireless Networks, 8*(5), 521–534.
57. Rajagopalan, S., Sankar, L., Mohajer, S., & Poor, H. (2011). Smart meter privacy: A utility-privacy framework. In *Proceedings of IEEE international conference on smart grid communications (SmartGridComm)* (pp. 190–195).
58. Gorman, S. (2009) Electricity grid in U.S. Penetrated by spies. *The Wall Street Journal*.
59. Margolis, J. (2010). From telecom privacy to utility privacy-coping with the needs of law enforcement on smart grid systems. Available at: http://www.narucmeetings.org/Presentations/100214%20Neustar%20NARUC%20Presentation.pdf (last accessed 2 Oct 2012).
60. National Institute of Standards and Technology. (2010). *Guidelines for Smart grid cyber security: Privacy and the smart grid*, Vol. 2. Cyber Security Working Group, NISTIR 7628, August 2010. Available at: http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol2.pdf (last accessed 10 Nov 2012).
61. U.S. Department of Energy. (2009). *DOE awards $620 million for ARRA 'Smart Grid' pilot projects*, December 10, 2009. Available at: bx.businessweek.com/smart-grid/ (last accessed 2 Oct 2012).
62. Fujioka, A., Okamoto, T., & Ohta, K. (1992). A practical secret voting scheme for large scale elections. In *Proceedings of AUSCRYPT '92: Workshop on the theory and application of cryptographic techniques*, LNCS 718, Queensland, Australia, December 1992 (pp. 244–251).
63. Benaloh, J. (1987). *Verifiable secret-ballot elections*. PhD thesis, Yale University.
64. Reiter, M., & Rubin, A. (1999). Anonymous Web transactions with crowds. *Communications of the ACM, 42*(2), 32–48.
65. Kamat P., Zhang Y., Trappe W., & Ozturk C. (2005). Enhancing source location privacy in sensor network routing. In *Proceedings of 25th IEEE international conference on distributed computing systems (ICDCS)*, Columbus, OH.
66. Jian, Y., Chen, S., Zhang, Z., & Zhang, L. (2007). Protecting receiver-location privacy in wireless sensor networks. In *Proceedings of Infocom 2007*, Anchorage, AK (pp. 1955–1963).
67. Ma, Z., Manglery, J., Wagner, C., & Bleier, T. (2011). Enhance data privacy in service compositions through a privacy proxy. In *Proceedings of ARES 2011*, Vienna, Austria (pp. 615–620).
68. Marmol, F., Sorge, C., Ugus, O., & Perez, G. (2012). Do not snoop my habits: Preserving privacy in the smart grid. *IEEE Communications Magazine, 50*(5), 166–172.
69. Fhom, H., & Bayarou, K. (2011). Towards a holistic privacy engineering approach for smart grid systems. In *Proceedings of the* 10th *IEEE international conference on trust, security, and privacy in computing and communications (TrustCom'11)*, Changsha, China (pp. 234–241).

Author Biographies

Sherali Zeadally is an Associate Professor in the Department of Computer Science and Information Technology at the University of the District of Columbia, Washington D.C., USA. He received his Bachelor's degree and Doctorate degree, both in Computer Science, from the University of Cambridge, UK and the University of Buckingham, UK respectively. He is a Fellow of the British Computer Society and a Fellow of the Institution of Engineering Technology, UK.

Al-Sakib Khan Pathan received Ph.D. degree in Computer Engineering in 2009 from Kyung Hee University, South Korea. He received B.Sc. degree in Computer Science and Information Technology from Islamic University of Technology (IUT), Bangladesh in 2003. He is currently an Assistant Professor at Computer Science department in International Islamic University Malaysia (IIUM), Malaysia and the Head, NDC Lab., KICT, IIUM. His research interest includes wireless sensor networks, network security, and e-services technologies. He is actively involved in various research activities and associated with various reputed journals and conferences as Editor, Chair, TPC member, and Reviewer.

Cristina Alcaraz received her PhD in Computer Science in 2011 from the University of Malaga and her M.Sc. in Computer Science degree in 2006. In 2007, she received her Master in Software Engineering and Artificial Intelligence. Her research activities are mainly focused on Critical Information Infrastructure Protection, and more precisely on secure monitoring of critical infrastructures, security of SCADA systems and Smart Grids, as well as the use of Wireless Sensor Networks for protection of critical systems. She is the author of several peer-reviewed journal and conference papers. She serves as a program committee member of several international conferences and as a reviewer of several international journals such as IEEE Transaction on Smart Grid, Sensors or Communications.

Mohamad Badra is currently employed by the College of Applied Science - Ministry of Higher Education - Sohar, Oman. His research interests include key exchange, wireless network security, public key infrastructures, smart cards, and wireless sensors networks. He received a PhD degree in networks and computer science from ENST-Paris in 2004. He is the author of several international standards on security exchange and the co-author of many international conference and journal papers.