The role of Signal Processing in Meeting Privacy Challenges [an overview]

Lalitha Sankar, Wade Trappe, Kannan Ramachandran, Harold Vincent Poor, Mérouane Debbah

The Role of Signal Processing in meeting Cyber-Security and Privacy Challenges

Lalitha Sankar, Wade Trappe, Kannan Ramachandran, H. Vincent Poor, and Mérouane Debbah

I. Introduction: Information Leakage Everywhere

We are in an era where computing and communication technologies are an integral part of our lives. Our environment is becoming increasingly more cyberphysical, with sensors and actuators exchanging information through a variety of networks. We access information from the cloud, share information with each other on social networks, and synthesize new information all while we are on-the-go. Storage technologies have also evolved significantly, with data density and access speeds projected to continue to increase dramatically over the next decade. The confluence of these technologies is leading to huge volumes of data flowing across our networks, to and from computers where they will be analyzed, and deposited into vast databases for later retrieval. The huge quantities of data poses new societal risks that are just now starting to manifest. Users post information on social networks unaware of the privacy risks they face; companies submit information to the cloud for processing to reduce their own computing costs yet unaware that potentially sensitive information might be leaked; sensor networks and global positioning systems monitor the state of our roads and vehicles yet reveal where we are driving to/from and even why.

Although some of these systems might be protected using conventional cryptographic protocols, many of them cannot and many of the risks exist outside the realm of protection offered by a cipher suite. For example, once information has been decrypted and placed into a database in plaintext it loses the confidentiality provided by the secure socket that connected the originating client to the database. Now, not only does the owner of that database know the exact value of that data, but it can examine the entirety of data it holds in order to glean information beyond the scope of the information provided.

To protect data from unforeseen security and privacy breaches, we should revisit information security and privacy from a fundamental point of view and search for techniques that complement traditional cryptographic services. Although conventional cryptographic protocols and services can and must be a part of the solution to ensuring data security and privacy, such techniques do not reflect the full spectrum of techniques available to protecting information. In particular, information is associated with a particular

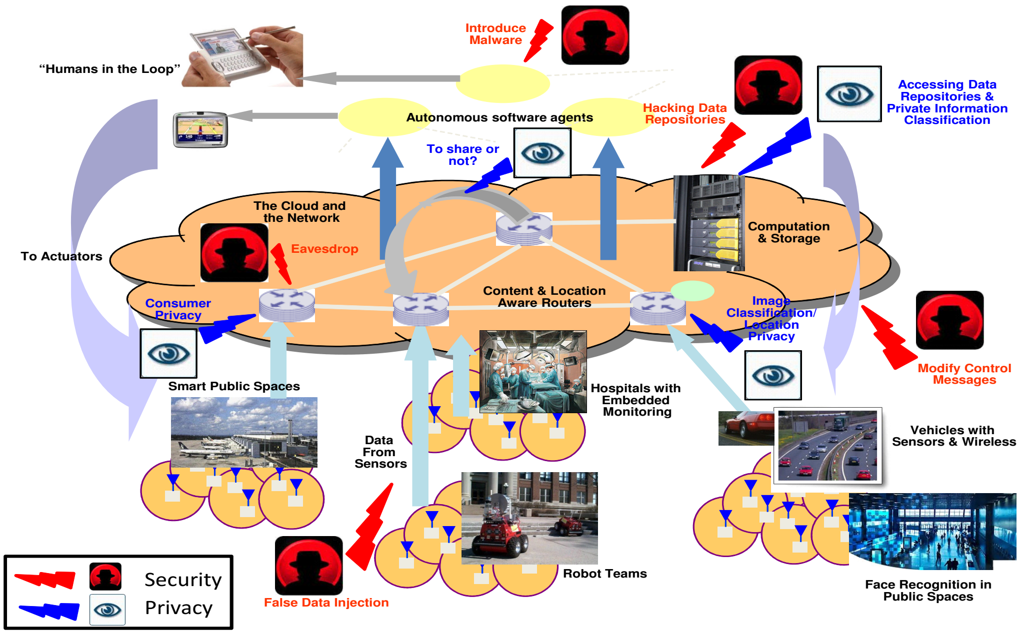

Fig. 1. Illustration of several classes of cyber-security and privacy threats in networks and distributed data systems.

context and has value to its users, and these notions of context and utility serve as the basis for formulating signal and information processing techniques that can enhance or complement traditional security/privacy services.

Whether one considers traditional network and computer security mechanisms, or complementary techniques for data protection that employ signal and information-theoretic mechanisms, there are several common classes of security threats that one expects to encounter. For example, in network communications an eavesdropper might listen on communications to gain access to restricted information, or the adversary might manipulate specific data fields to falsify data or instructions. Similarly, in cloud storage applications the adversary might be the storage service itself and attempt to glean unintended information from its repositories that was meant to remain private, or a user might inject false information into the database repository in hopes of harming the utility of the database to others as illustrated in Figure 1. At a high-level there are several major goals that can help address in order to better protect information. Three of these objectives, which correspond to the traditional “C-I-A” security framework, are described below, where we have outlined how signal processing techniques can have a direct impact.

- *Confidentiality:* Confidentiality is a guarantee that only those entities intended to access information or data should be able to access that information. In communications, this might involve a guarantee that the information exchanged in a protocol is only able decipherable by legitimate entities. In database applications, where the data from many users might be stored together, confidentiality involves encrypting the data and managing the keys needed to access and decrypt the data.

- *Integrity:* Integrity is a guarantee that a system or data is consistent with its purpose or how it is expected to be shared. For example, in communications, integrity might involve algorithms (e.g., cryptographic checksums) that ensure that data or messages have not been altered, or provide assurance that the data came from its claimed origin. In network and computer systems, integrity corresponds to ensuring that the system is operating correctly and that anomalous events can be rapidly detected and defended against. Similarly, in database and data-driven applications, it is desirable to ensure that stored or actuated upon data is correct and consistent with physical reality.
- *Availability:* The goal of availability is to ensure that system resources, such as stored data or services, are accessible to entities when legitimate service requests are made. In the context of communications and network services, this might correspond to ensuring protecting against denial of service attacks, while in computer systems this involves guaranteeing that if an entity has privileges for specific set of objects/services then that access should not be obstructed.

Confidentiality as a security goal is particularly tricky to categorize since it can involve issues such as preventing an eavesdropper from learning specific information, ensuring that various contexts surrounding data is preserved, and that the combination of several distinct pieces of data, when put together, do not allow an adversary to infer other information that was not meant to be inferred. For the sake of clarity, in this paper we shall specifically call out two different flavors of confidentiality: *secrecy*, which is concerned with ensuring that specific information cannot be received/deciphered by well-defined and untrusted adversaries external to the system (e.g. eavesdroppers), and *privacy* which involves ensuring that unintended information is not revealed (via inference and correlations) when intended data is shared with legitimate users.

In this article, we shall examine signal and information processing techniques that can be used to protect information. We begin in Section II by identifying several information-theoretic (IT) themes that have emerged across many emerging cyber-security and privacy problems. After identifying these basic themes, we then explore in detail how signal processing techniques that are motivated by IT principles can be applied to address emerging privacy problems.

II. An Information-Processing Perspective of Security

Security and privacy, at a high-level, involves making certain that an adversary does not know the correct value or meaning behind information being shared and that the adversary cannot alter or affect the data or a system in a manner that is not allowed by legitimate parties.

Let us generically label an original piece of data as $\mathbf{X}$, which may correspond to a communications signal or to N different fields in a database or to an N-dimensional data vector measured by a cyber-physical sensor system. The confidentiality objective, whether secrecy or privacy, involves applying a transformation on $\mathbf{X}$ to create a new information $\mathbf{Y}$ vector from which it is difficult for an adversary to answer questions about $\mathbf{X}$ or even infer $\mathbf{X}$ itself. For example, in conventional cryptography, $\mathbf{Y}$ corresponds to ciphertext that is the result of an encryption algorithm being applied to $\mathbf{X}$. In general, some form of advantage should exist between the legitimate parties (Alice and Bob) over the adversary (Eve). For conventional cryptography, this advantage takes the form of Alice and Bob sharing an encryption key unknown to Eve. In many settings, it is possible to find other sources of advantage that Alice and Bob share over Eve. For example, there has been recent work in the physical layer security community, where the properties of the (wireless) communications is exploited for secrecy. This area of research is well-grounded in IT principles [1], [2], and the body of work suggests an intuitive strategy for achieving secret communications: one should develop methods that ensure that the Alice-Bob channel effectively has higher quality than the Alice-Eve channel.

However, since many of these methods are analogous to mechanisms employed in privacy, we briefly give our own interpretation of signal and information processing strategies that may be used to enhance confidentiality. Later, in Section III, we commence with exploring such methods in more detail in terms of privacy where the goal is to ensure that unintended information is not revealed to legitimate users.

- *Perturbing Data:* A natural approach to protecting the true data value is to introduce a perturbation to the data by adding random noise. Such an approach has been applied, for example, in the database privacy community [3], where one attempts to design a noise distribution so as to obscure the original data distribution while preserving certain aggregate metrics (e.g. data average). Another area where perturbation is applied is in secret communications, where legitimate parties design interference signals so as to best disrupt an eavesdropper's ability to listen on communications [4]. In both cases, confidentiality becomes a problem in signal design and signal processing techniques can be applied.
- *Altering Data Precision:* Data has utility that is directly related to its precision. In many situations, it is possible to exploit the different valuations that legitimate users have in the data. For example,

a legitimate user might only need coarse-grained information, while an adversary might only derive value and infer on fine-grained information. In such scenarios, it is possible to quantize the data to a level coarser than originally measured [5]. Such techniques arise in problems like location privacy, where coarse-grained position information is released to a location-service.

- *Data Aggregation:* Another approach is to amass a collection of information and, rather than report all of the data, merely report an aggregation of that data. For example, a location-based service might choose to reveal that a group of users is at a specific location (e.g. a specific office) or that an individual is at a vague location (e.g. a shopping mall as opposed to a specific store), but would not reveal that a specific individual is at a specific location [5]. Similar techniques have been employed in Web services, such as in [6], where users are gathered into geographically diverse groups to make it difficult to identify who is initiating a Web request. In all of these cases, techniques from signal and information fusion can be used to formulate the tradeoff between privacy and data utility.

Moving beyond confidentiality, integrity also involves examining $\mathbf{X}$ to decide whether $\mathbf{X}$ corresponds to a valid situation. For example, we might want to decide whether temperature data from a thermistor used in a smart home is correct. It is clearly unreasonable to accept a temperature reading of 10000 Kelvin. Similarly, in network anomaly detection, $\mathbf{X}$ might correspond to traffic statistics measured over a period of time and we might wish to infer whether $\mathbf{X}$ corresponds to an acceptable operational condition for the network, or whether $\mathbf{X}$ flags anomalous activities [7]. As a final example, we might have data being inserted by a malicious entity into a database or into cloud storage that will be accessed by a wide array of users. In such situations, it is often desirable for the database service to sanitize the database of false or outlier data. In any of these cases, a general signal or information processing approach to assuring the integrity of information or a system involves formulating the integrity problem using a statistical hypothesis testing framework. Specifically, hypothesis checking involves determining whether $\mathbf{X}$ falls inside or outside of a region of validity Ω. If the data falls in the valid region, the null hypothesis H_0 is verified and thus the data/system integrity is validated. On the other hand, unreliable data is classified as H_1, and the system integrity is deemed suspect or compromised.

III. Protecting Private Information

While security mechanisms address the problem of external hackers and eavesdroppers, guaranteeing the *privacy* of individuals and organizations whose data is being transferred across networks and stored in distributed repositories requires different tools to protect contextual information from being gleaned by potential *insider* inference.

There are many application scenarios where the issue of privacy arises and unfortunately the traditional approach to developing privacy solutions for these applications involves application-specific solutions that are often heuristic and generally disconnected from a formal theoretical foundation that quantifies the privacy benefits. These different applications, whether they correspond to smart grids or medical records, can all benefit from a unifying framework that allows a privacy engineer to build solutions where the privacy benefits can be explored with well-understood utility tradeoffs, first introduced and developed in . In this section, we begin by outlining a unified approach to understanding privacy and application utility. This model is based on IT principles introduced and developed in [12]. Following the model presentation, in Section IV we illustrate how this model can be used to motivate the application of signal processing techniques to reliably support information privacy.

A. A Unifying Analytical Model

Our privacy model involves datasets (e.g. databases or communication packets/signals) with K attributes per entry/occurrence. For the remainder of this discussion, we shall simplify our discussion to refer to these datasets as databases. Our proposed model focuses on large databases with K attributes per entry. Let $\mathcal{X}_k$, for all $k \in \mathcal{K} = \{1, 2, , \ldots, K\}$, and $\mathcal{Z}$ be finite sets. Let $X_k \in \mathcal{X}_k$ be a random variable denoting the k^{th} attribute, $k = 1, 2, \ldots, K$, and let $X_{\mathcal{K}} \equiv (X_1, X_2, \ldots, X_K)$. A database d with n rows is a sequence of n independent observations from a source having a probability distribution

$$p_{X_{\mathcal{K}}}(x_{\mathcal{K}}) = p_{X_1 X_2 \ldots X_K}(x_1, x_2, \ldots, x_K) \tag{1}$$

which is assumed to be known to both the designers and users of the database. Our simplifying assumption of row independence holds generally in large databases as correlation typically arises across attributes and can be ignored across entries given the size of the database. We write $X_{\mathcal{K}}^n = (X_1^n, X_2^n, \ldots, X_K^n)$ to denote the n independent and identically distributed (i.i.d.) observations of $X_{\mathcal{K}}$.

The joint distribution in (1) models the fact that the attributes corresponding to an individual entry are correlated in general and consequently can reveal information about one another.

Public and private attributes: We consider a general model in which some attributes need to be kept private while the source can reveal a function of some or all of the attributes. We write $\mathcal{K}_r$ and $\mathcal{K}_h$ to denote sets of private (subscript h for hidden) and public (subscript r for revealed) attributes, respectively, such that $\mathcal{K}_r \cup \mathcal{K}_h = \mathcal{K} \equiv \{1, 2, \ldots, K\}$. We further denote the corresponding collections of public and private attributes by $X_{\mathcal{K}_r} \equiv \{X_k\}_{k \in \mathcal{K}_r}$ and $X_{\mathcal{K}_h} \equiv \{X_k\}_{k \in \mathcal{K}_h}$, respectively.

Our notation allows for an attribute to be both public and private; this is to account for the fact that a database may need to reveal a function of an attribute while keeping the attribute itself private. In general, a database can choose to keep public (or private) one or more attributes ($K > 1$). Irrespective of the number of private attributes, a non-zero utility results only when the database reveals an appropriate function of some or all of its attributes.

Revealed attributes and side information: The public attributes $X_{\mathcal{K}_r}$ are in general sanitized/distorted prior to being revealed in order to reduce possible inferences about the private attributes. We denote the resulting *revealed attributes* as $\hat{X}_{\mathcal{K}_r} \equiv \{\hat{X}_k\}_{k\in\mathcal{K}_r}$. In addition to the revealed information, a database user can have access to correlated side information from other information sources. We model the side information as an n-length sequence $Z^n = (Z_1, Z_2, \ldots, Z_n)$, $Z_i \in \mathcal{Z}$ for all i, which is correlated with the database entries via a joint distribution $p_{X_{\mathcal{K}}Z}(x_{\mathcal{K}},z)$.

Reconstructed database: The final *reconstructed database* at the user will be either a database of revealed public attributes (when no s.i. is available) or a database generated from a combination of the revealed public attributes and the side information (when s.i. is available).

B. The Privacy and Utility Principle

Even though utility and privacy measures tend to be specific to the application, there is a fundamental principle that unifies all these measures in the abstract domain. The aim of a privacy-preserving database is to provide some measure of utility to the user while at the same time guaranteeing a measure of privacy for the entries in the database.

A user perceives the utility of a perturbed database to be high as long as the response is similar to the response of the unperturbed database; thus, the utility is highest of an unperturbed database and goes to zero when the perturbed database is completely unrelated to the original database. Accordingly, our utility metric is an appropriately chosen average 'distance' function between the original and the perturbed databases. Privacy, on the other hand, is maximized when the perturbed response is completely independent of the data. Our privacy metric measures the difficulty of extracting any private information from the response, i.e., the amount of uncertainty or *equivocation* about the private attributes given the response [14], [12]. This basic principle and the resulting mappings from the application to the analysis (abstract) domain are summarized in Table I.

C. A Privacy-Utility Tradeoff Model

Since database sanitization is traditionally the process of distorting the data to achieve some measure of privacy, it is a problem of mapping a database to a different one subject to specific utility and privacy requirements.

Application Requirements	Principle/Abstraction
Utility	Distortion/Fidelity Functions
Privacy	Equivocation
Perturbation Technique	Information-theoretic Source Code

TABLE I

A PRINCIPLED MAPPING OF APPLICATION REQUIREMENTS TO AN ABSTRACT MODEL.

Mapping: Our notation below relies on this abstraction [12]. Let $\mathcal{X}_k, k \in \mathcal{K}$, and $\mathcal{Z}$, be as above and let $\hat{\mathcal{X}}_j$ be additional finite sets for all $j \in \mathcal{K}_r$. Recall that a database d with n rows is an instantiation of $X^n_{\mathcal{K}}$. Thus, we will henceforth refer to a real database d as an *input database* and to the corresponding sanitized database (SDB) d_s as an *output database*. When the user has access to side information, the *reconstructed database* d' at the user will in general be different from the output database.

Our coding scheme consists of an encoder F_E which is a mapping from the set of all input databases (i.e., all databases d allowable by the underlying distribution) to a set of indices $\mathcal{J} \equiv \{1, 2, \ldots, M\}$ and an associated table of output databases (each of which is a d_s and has a one-to-one mapping to $\mathcal{J}$) given by

$$F_E : (\mathcal{X}^n_1 \times \ldots \times \mathcal{X}^n_k)_{k \in \mathcal{K}_{enc}} \to \mathcal{J} \equiv \{SDB_k\}^M_{k=1} \tag{2}$$

where $\mathcal{K}_r \subseteq \mathcal{K}_{enc} \subseteq \mathcal{K}$ and M is the number of output (sanitized) databases created from the set of all input databases. To allow for the case where an attribute can be both public and private, we allow the encoding F_E in (2) to include both public and private attributes. A user with a view of the SDB (i.e., an index $j \in \mathcal{J}$) and with access to side information Z^n, whose entries Z_i, $i = 1, 2, \ldots, n$, take values in the alphabet $\mathcal{Z}$, reconstructs the database d' via the mapping

$$F_D : \mathcal{J} \times \mathcal{Z}^n \to \left(\prod_{k \in \mathcal{K}_r} \hat{\mathcal{X}}^n_k \right). \tag{3}$$

The encoding and decoding are assumed known at both parties.

Utility: Relying on a distance based utility principle, we model the utility u via the requirement that the average *distortion* of the public variables is upper bounded, for each $\epsilon > 0$ and all sufficiently large n, as

$$u \equiv \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n \rho\left(X_{\mathcal{K}_r,i}, \hat{X}_{\mathcal{K}_r,i}\right)\right] \leq D + \epsilon, \tag{4}$$

where $\rho(\cdot,\cdot)$ denotes a distortion function, $\mathbb{E}$ is the expectation over the joint distribution of $(X_{\mathcal{K}_r}, \hat{X}_{\mathcal{K}_r})$, and the subscript i denotes the i^{th} entry of the database. Examples of distortion functions include the

Euclidean distance for Gaussian distributions, the Hamming distance for binary input and output databases, and the Kullback-Leibler (K-L) divergence.

Privacy: We quantify the equivocation e of all the private variables using entropy as

$$e \equiv \frac{1}{n} H\left(X^n_{\mathcal{K}_h} | J, Z^n\right) \geq E - \epsilon. \tag{5}$$

Analogous to (5), for continuous sources, we can quantify the privacy leakage l using mutual information as

$$l \equiv \frac{1}{n} I\left(X^n_{\mathcal{K}_h}; J, Z^n\right) \leq L + \epsilon. \tag{6}$$

The mappings in (2) and (3) ensure that d is mapped to d' such that the constraints in (4) and (5) are met. The formalism in (1)-(6) is analogous to lossy compression in that a source database is mapped to one of M quantized databases that are designed *a priori*. For a chosen encoding, a database realization is mapped to the appropriate quantized database. It suffices to communicate the index J of the resulting quantized database as formalized in (2) to the user. This index, in conjunction with side information, if any, enables a reconstruction at the user as in (3). *Note that the mappings in (2) and (3), i.e., lossy compression with privacy guarantees, ensure that for any $D > 0$, the user can only reconstruct the database $d' = \hat{X}^n_{\mathcal{K}_r}$, formally a function $f(J, Z^n)$, and not $d = X^n_{\mathcal{K}}$ itself.*

The utility and privacy metrics in (4) and (5) capture the statistical nature of the problem, i.e., the fact that for large enough n the entries of the database statistically mirror the distribution (1). Thus, both metrics represent averages across all database instantiations d, and hence, (assuming stationarity and large n) over the sample space of $X_{\mathcal{K}}$ thereby quantifying the average distortion (utility) and equivocation (privacy) achievable per entry.

D. Equivalence of Utility-Privacy and Rate-Distortion-Equivocation

Mapping utility to distortion and privacy to information uncertainty via entropy (or leakage via mutual information) leads to the following definition of the utility-privacy (U-P) tradeoff region.

Definition 1: The U-P tradeoff region $\mathcal{T}$ is the set of all U-P tuples (D, E) for which there exists a coding scheme (F_E, F_D) given by (2) and (3), respectively, with parameters (n, M, u, e) satisfying (4) and (5).

While $\mathcal{T}$ in Definition 1 can be determined for specific database examples, one has to, in general, resort to numerical techniques to solve the optimization problem. To obtain closed form solutions that define the set of all tradeoff points and identify the optimal encoding schemes, we exploit the rich set of techniques from rate distortion theory with and without equivocation constraints. To this end, we study a more general problem of rate-distortion-equivocation (RDE) by introducing an additional rate constraint $M \leq 2^{n(R+\epsilon)}$ which bounds the number of quantized SDBs in (2). Besides enabling the use of known rate-distortion techniques, the rate constraint also has an operational significance. For a desired level of accuracy (utility) D, the rate R is the precision required on average (over $\mathcal{X}_K$) to achieve it. We now define the achievable RDE region as follows.

Definition 2: The RDE region $\mathcal{R}_{RDE}$ is the set of all tuples (R, D, E) for which there exists a coding scheme given by (2) and (3) with parameters (n, M, u, e) satisfying the constraints in (4), (5), and on the rate. In this region, $\mathcal{R}_{D-E}$, the set of all feasible distortion-equivocation tuples (D, E) is defined as

$$\mathcal{R}_{D-E} \equiv \{(D, E) : (R, D, E) \in \mathcal{R}_{RDE},\ R \geq 0\}\,. \tag{7}$$

The RDE problem includes a constraint on the precision of the public variables in addition to the equivocation constraint on the private data in the U-P problem. Thus, in the RDE problem, for a desired utility D, one obtains the set of all rate-equivocation tradeoff points (R, E), and therefore, over all distortion choices, the resulting region contains the set of all (D, E) pairs. From Definitions 1 and 2, we thus have the following Proposition.

Proposition 3 ([12]): $\mathcal{T} = \mathcal{R}_{D-E}$.

Proposition 3 is captured pictorially in Fig. 2(b). The functions $R\,(D, E)$ and $\Gamma(D)$ in Fig. 2 capture the rate and privacy boundaries of the region and are the minimal rate and maximal privacy achievable, respectively, for a given distortion D. Furthermore, in Fig. 2(b) the U-P tradeoff region contrast existing privacy-exclusive and utility-exclusive regimes (extreme points of the utility-privacy tradeoff curve) with our general approach of determining the set of all feasible utility-privacy tradeoff points.

The aim of classical source coding is to determine the smallest number of bits per sample, i.e., the rate-distortion function $R(D)$, that when revealed to the user (decoder), ensures reconstruction within a distortion D. When an additional privacy constraint in (5) is included, the problem becomes one of determining the achievable RDE region. Thus, a RDE code is by definition a (lossy) source code (in practice a compression algorithm) satisfying a distortion constraint that achieves a specific privacy level for every choice of distortion. The power of Proposition 3 is that it allows us to study the larger problem of database U-P tradeoffs in terms of a relatively familiar problem of source coding with additional privacy constraints. This result shows the tradeoff between utility (distortion), privacy (equivocation), and precision (rate) – fixing the value of any one determines the set of operating points for the other two; for example, fixing the utility (distortion D) quantifies the set of all achievable privacy-precision tuples (E, R).

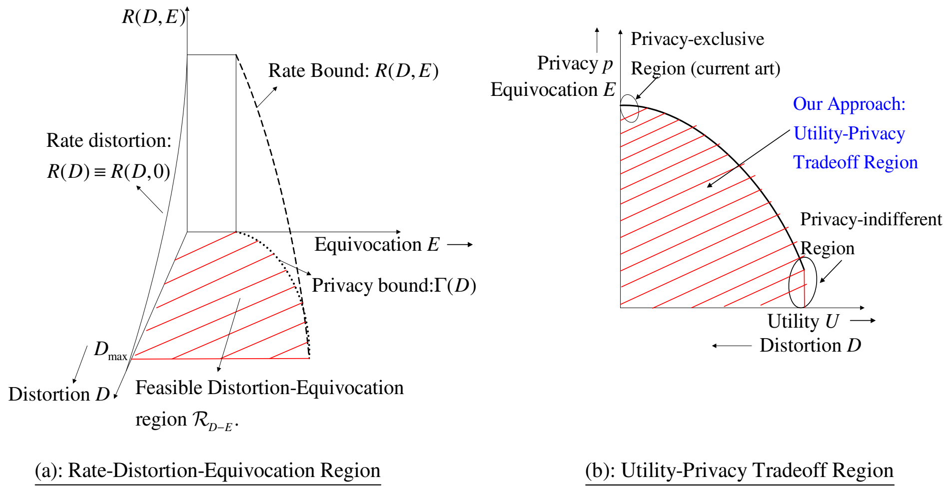

Fig. 2. (a) Rate Distortion Equivocation Region; (b) Utility-Privacy Tradeoff Region.

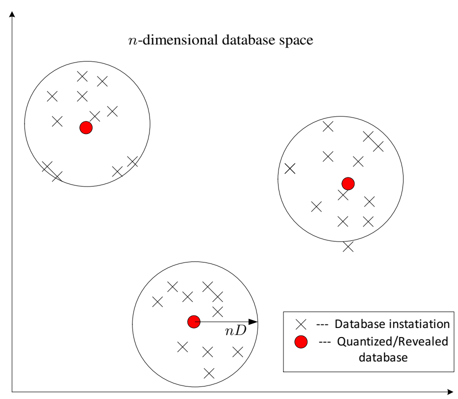

Fig. 3. Space of all database realizations and the quantized databases.

E. Understanding the Formalism

The formalism here aims to determine the optimal sanitization, i.e., a mapping which guarantees the maximal privacy for the private (hidden) attributes and a desired level of utility for the public attributes, among the set of *all* possible mappings that transform the public attributes of a database. We use the terms *encoding* and *decoding* to denote this mapping at the data publisher end and the user end respectively. A database instance is an n-realization of a random source (the source is a vector when the number of attributes $K > 1$) and can be viewed as a point in an n-dimensional space (see Fig. 3). The set of all possible databases (n-length source sequences) that can be generated using the source statistics (probability distribution) lie in this space.

Our choice of utility metric is a measure of average 'closeness' between the original and revealed database public attributes via a distortion requirement D. Thus the output of sanitization will be another database (another point in the same n-dimensional space) within a ball of 'distance' nD. We seek to determine a set of $M = 2^{nR}$ output databases that 'cover' the space, i.e., given any input database instance there exists at least one sanitized database within bounded 'distance' nD as shown in Fig. 3. Note that the sanitized database may be in a subspace of the entire space because only the public attributes are sanitized and the utility requirement is only in this subspace.

Such a distortion-constrained encoding is referred to in the signal processing literature as vector quantization because the compression is of an n-dimensional space and can be achieved in practice using clustering algorithms. In addition to a distortion (utility) constraint, our privacy constraint also requires that the "leakage" (i.e. the loss of uncertainty) about the private attributes via correlation from the sanitized database is bounded. The set of M source-sanitized database pairs is chosen to satisfy both distortion and leakage constraints.

IV. Privacy Case Studies

We now turn our attention to exploring case studies where signal and information processing techniques, motivated by the framework we outline in Section III, can be applied to improve privacy. Our exploration of case studies will examine three broad categories of database privacy: (i) *statistical data privacy* which involves guaranteeing privacy of any individual in a database that is used for statistical information processing (utility); (ii) *competitive privacy* which involves information sharing for a common system good (utility) between competing agents that are physically interconnected; and (iii) *consumer privacy* related to guaranteeing privacy when monitoring users using smart devices (utility).

A. Statistical Data Privacy

The problem of privacy for databases was first exposed by census statisticians who were required to publish statistics related to census functions but without revealing any particulars of individuals in the census databases [9]. Several early attempts were made to publish census data using ad hoc techniques such as sub-sampling. However, the first widely reported attempt at a formal definition of privacy was by Sweeney [13]. The concept of k-*anonymity* proposed by Sweeney captures the intuitive notion of privacy that every individual entry should be indistinguishable from $(k-1)$ other entries for some large value of k.

The approaches considered in the literature have centered on the correct application of *sanitization* that ensure that a user interacts only with a modified database that is derived from the original (e.g.: [9], [13], [3]). Most of the these sanitization approaches, with the exception of differential privacy-based ones, are application-specific and often focus on additive noise approaches. More recently, privacy approaches for statistical databases has been driven by the differential privacy definition [11] in which the authors take the view that the privacy of an individual in a database is related to the ability of an adversary to detect whether that individual's data is in that database or not. In this special issue, the reader will find an article *****To Be Filled in After Review Cycle Completes***** discussing signal processing techniques for DP.

DP is a strictly stronger concept than the IT definition of privacy used here. However, the IT model has a strong statistical basis and is suited to many application domains where strict anonymity is not the requirement. For example, in many wellness databases the presence of the record of an individual is not a secret but that individual's disease status is. The vector quantization based sanitization approach presented here applies to both numerical and categorical data whereas DP, while being a very popular model for privacy, appears limited to numerical data.

More generally, a rigorous model for privacy-utility tradeoffs with a method to achieve *all* the optimal points has remained open and is the subject of this survey. The use of information theoretic and signal processing tools for privacy and related problems is relatively sparse and and is beginning to draw the attention of signal processing researchers. [14] analyzed a simple two variable model using rate distortion theory with equivocation constraints, which is the prime motivation for the general framework introduced here.

We now illustrate how our framework can be applied to a simple numerical database example. Consider a $K = 2$ database where both attributes X and Y are jointly Gaussian with zero means and variances σ_X^2 and σ_Y^2, respectively, and with correlation coefficient $\rho = E\left[XY\right] / \left(\sigma_X \sigma_Y\right)$. This model applies for numeric data (e.g., clinical data), such as height and weight measures, which are roughly assumed to be normally distributed. We assume that for every entry only one of the two attributes, say X, is revealed while the other, say Y, is hidden such that $Y - X - \hat{X}$ forms a Markov chain. The RDE region for this case can be obtained directly from Yamamoto's results [14] with appropriate substitution for a jointly Gaussian source. Furthermore, due to the Markov relationship between of X, Y, and $\hat{X}$, the minimization of $I(X;\hat{X})$ is strictly over $p(\hat{x}|x)$, and thus, simplifies to the familiar rate-distortion problem for a Gaussian source X which in turn is achieved by choosing the reverse channel from $\hat{X}$ to X as an additive white Gaussian noise channel with variance D (average distortion). The resulting minimal rate

$R(D)$ and leakage $L(D)$ achieved (in bits per entry) are, for $D \in \left[0, \sigma_X^2\right]$, $R(D) = \frac{1}{2}\log\left(\sigma_X^2/D\right)$ and $L(D) = \frac{1}{2}\log(\left[\left(1-\rho_{XY}^2\right) + \rho_{XY}^2 D \left/\sigma_X^2\right.\right]^{-1})$, respectively. As expected, $L(D)$ is maximal for $D = 0$ (X revealed perfectly) in which case only the data independent of X in Y can be guaranteed perfect privacy, and minimal for $D = \sigma_X^2$ when nothing is revealed ($R = 0$). The largest U-P tradeoff region is thus the region enclosed by $L(D)$.

It is worth noting that the reconstructed database is also Gaussian distributed. Thus, if one wishes to preserve the most uncertainty about (some or all of) the original input database from the output, it suffices to add Gaussian noise. The power of the framework are two-fold: (i) one can find the privacy-optimal sanitization for the Gaussian case; and (ii) practical applications such as medical analytics that assume Gaussian-distributed data can still work on sanitized data, albeit with modified parameter values.

B. Competitive Privacy

The problem of sharing information across autonomous entities in such a way that no information apart from the answer to the query is revealed was identified by Agrawal *et al* [16] (see also [17] where the term competitive privacy is introduced). More recently, in [18], the notion of competitive privacy was generalized to a variety of domains involving complex distributed networks (e.g., financial, water distribution, transport and congestion control, etc.).

The evolution of distributed sensing, control, and actuation systems to monitor complex networks (e.g., the power grid) had led to the problem of interactive distributed processing, often between competing entities, that manage subnetworks of a large connected network. The common benefit of network reliability drives the need for interaction while economics, competition, and consumer policies lead to privacy (confidentiality) constraints. In [18] the formalism presented here is applied to the problem of distributed state estimation (estimation of complex voltages across the grid) enabled via data interactions amongst many agents. The formalism captures the tradeoff at each agent between the rate (precision) of data exchanges, the fidelity of its estimate (distortion), and the leakage of its private information (typically state information, say X, in an agent's subnetwork) when every agent shares a (compressed) function of its measurements (say Y which is a function of some or all X) with others.

For a Gaussian measurement model (i.e., measurement Y at each agent is a noisy linear combination of Gaussian states X at the different agents), [18] shows that while private information leakages are inevitable in an interactive model, it is possible to share data minimally to achieve both the desired global estimation reliability at each agent and minial information leakage. More broadly, the resulting RDE tradeoff framework demonstrates that privacy- and confidentiality-aware interactions are possible in a number of distributed computing applications.

Privacy and confidentiality concerns are frequently driven by economics [20]; for the distributed state estimation problem described above, [19] explores the *cost of privacy* using game-theoretic principles and shows that strictly non-zero prices are required to incentive cooperation when agents choose to optimize a linear combination of their leakage and distortion functions. For the same setup, [21] highlights the power of a repeated games model to build trust and enable collaborate over multiple interactions. Signal processing algorithms that allow interaction and games is the first step towards enabling secure distributed platforms.

C. Consumer Privacy

We now examine consumer privacy, which deals with problems arising as organizations monitor their clients or users. Within this broad class of client privacy, there exist different types of information that can be collected. One such category consists of physical measurements; this includes quantities like water, energy, or bandwidth usage. For instance, when people use electricity, power companies track clients' usage patterns in order to determine pricing; on the other hand, detailed information about power usage can reveal a lot about an individual's daily habits. The client would ideally like to preserve some level of privacy while still being able to benefit from the power company's service. It thus makes sense to ask how much information really needs to be leaked without affecting service quality. We illustrate this category with an example of privacy for smart meters and highlight the signal processing challenges.

Another category consists of more categorical measurements, such as the types of products a person buys, or the nature of Internet queries. In this context, we will discuss an example of how signal processing techniques can be used to achieve privacy in an image classification application.

1) Smart Meter Privacy: While the fine-grained monitoring afforded by smart meters enable better load balancing and reliability analysis in the electric grid, data from smart meters may be mined to infer personal information in ways that are unknown to us presently. A number of solutions have been proposed such as using batteries to mask usage patterns [22], [23], anonymizing user data, adding noise and aggregating neighborhood data [24]. Focusing on the tradeoff problem, [25] proposes a rigorous approach grounded in the framework presented in the earlier Section to quantify the tradeoff between utility and privacy and reveal signal processing methods of achieving privacy.

The motivation is to decouple the revealed meter data as much as possible from the personal actions of a consumer. This insight is based on the observation that irregular (intermittent) activity such as kettles or lights turned on manually are much more revealing of personal actions than regular (continuous) activity such as refrigerators or lights on timers. In [25], the time-series data from the smart meters is modeled as a noisy linear mixture of colored Gaussian processes[1], each process generated by an underlying appliance. Appliances are modeled as belonging to two broad classes, one which is continuously on (e.g., air conditioners, heaters) in time (and thus, have narrow spectral signatures) and the other which come on intermittently (e.g., toasters, kettles, washers, etc.) and are bursty in time (and thus, have broad spectral signatures). The privacy-utility tradeoffs on the total load are shown to be achievable using an *interference-aware reverse waterfilling* solution which: (i) exploits the presence of high-power but less private appliance spectra as implicit distortion noise, and (ii) filters out frequency components with lower power relative to a distortion threshold.

The solution in [25] exploits the fact that the time-series measurements can be transformed to the spectral domain in which minimizing the information leakage of the intermiittent appliances leads to the problem of optimally allocating the distortion (mean-squared error between the actual and revealed meter data) for each spectral frequency subject to a joint distortion constraint over all frequencies. This joint distortion constraint in turn determines a *threshold waterlevel* (PSD) below which all spectral energy is suppressed (and hence, the name *reverse water-filling*).

The power of the solution is illustrated in the following example. A continuous (denoted c) and an intermittent appliance (denoted i) load measurements are modeled as (time-limited) Gauss-Markov processes whose auto-correlation fucntions $P_{(l)}\rho_{(l)}^{-|k|}$, $l = i, c$, are characterized by variance $P_{(l)}$, correlation coefficient $\rho_{(l)}$, and memory $m_{(l)}$ for the l^{th} appliance type. These parameters allow modeling the fact that the continuously used appliances have a longer memory and a larger correlation coefficient relative to the intermittently used appliances; furthermore, the bursty usage pattern of the intermittent appliances is incorporated by choosing $P_i > P_c$.

For $P_i = 12$, $P_c = 8$, $\rho_i = 0.4$, $\rho_c = 0.8$, $m_i = 40$, and $m_c = 120$, in Fig. 4, the power spectral densities (PSDs) for the intermittent, continuous (with noise), and both (with noise) for the parameters is shown. Also shown is the waterlevel λ and the distortion spectrum $\Delta(f)$. From both figures, we see that the distortion spectrum is zero when the PSD of the noisy continuous process dominates either that of the intermittent process or the waterlevel λ leading to zero and minimal leakage, respectively, for the two cases. The waterlevel λ determines the distortion spectrum otherwise.

2) Image Classification Privacy: A commonly-cited application for client privacy is biometric identification [26]. The FBI recently unveiled plans to spend one billion dollars on a tracking system that

[1]Colored processes are those in which the signals generated at each time instance are correlated with those generated previously.

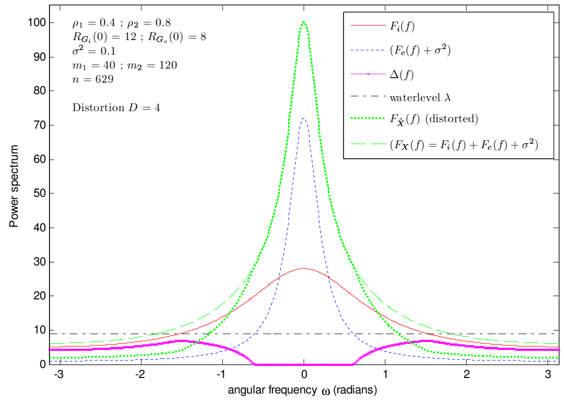

Fig. 4. A graphical illustration of the spectral filtering approach to suppress those appliance signatures that if revealed can lead to privacy violations.

relies on face recognition to scan surveillance video for wanted persons; meanwhile, civil rights groups are raising objections about privacy violations [27]. To relieve this tension, privacy-preserving algorithms could prevent normal citizens from being tracked by the government, while simultaneously allowing police agencies to find criminals. This problem reduces to one of privacy-preserving image classification. Many existing algorithms dealing with private media classification rely on cryptographic primitives that permit basic computation in the encrypted domain (e.g., [28]). An example of a system that instead uses signal processing techniques to reduce the necessary communication and computational load is the private face recognition system in [29]. In this system, the client inputs a face image, and with the help of a server storing a database of faces, the client learns the subject's name. Meanwhile, the server learns nothing about the client's query.

The basic insight in [29] is that a practical privacy-preserving search algorithm should be designed with privacy primitives in mind. Therefore, the system alters a Euclidean-distance nearest-neighbor search into a Hamming distance comparison, which is more compatible with distortion-intolerant privacy primitives. This problem redefinition occurs via a feature vector transformation that involves projection onto random hyperplanes [30]. Yeo et al. showed that the Hamming distance between two transformed vectors is directly related to the Euclidean distance between the original real-valued vectors. The new binary features couple nicely with a primitive called private information retrieval to enable privacy-preserving searches at communication and computation costs sublinear in database size.

This example is useful because it illustrates that signal processing techniques—in this case, the feature transformation—can significantly reduce the resource consumption of privacy-aware systems. However, the face recognition system in [29] combines signal processing with privacy primitives to achieve IT client privacy. An even more signal-processing-flavored option would be a system that eschews privacy primitives altogether, and instead relies on the privacy properties of transformed feature vectors. Of course, to make such a system useful, it is critical to understand the privacy properties of various feature transformations at a theoretical level.

V. Concluding Remarks

We have explored a set of non-cryptographic mechanisms for security and privacy, and provided a unifying framework for studying and formulating the tradeoff between privacy and data utility. We note that traditional security mechanisms must play a role in securing our communications and information. However, there are many situations on the horizon where a set of new tools can complement and enhance traditional security mechanisms. Domain-specific approaches that are founded on well-developed information-theoretical and signal processing principles represent an untapped resource for developing such new security tools. We explored several case studies where the methods described in this survey article can be used to drive the design of signal and information processing techniques that can thwart threats that cannot be dealt with using only conventional cryptographic mechanisms.

References

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. Journal*, vol. 54, pp. 1355–1387, 1975.

[2] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, pp. 339–348, 1978.

[3] D. Agrawal and C. Aggarwal, "On the design and quantification of privacy preserving data mining algorithms," in *Proc. 20th Symp. Prin. Database Systems*, Santa Barbara, CA, May 2001.

[4] Y. Liang, V. Poor and S. Shamai, "Secure communication over fading channels," *IEEE Trans. Info. Theory*, vol. 54, pp. 2470-2492, 2008.

[5] M. Gruteser and D. Grunwald, "Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking," in *Proc. 1st Intl. Conf. Mobile Syst., Apps. and Services (Mobisys'03)*, pp. 31-42, 2003.

[6] M. Reiter and A. Rubin, "Crowds: anonymity for web transactions," *ACM Trans. Inform. Syst. Security*, vol. 1, pp. 66–92, 1998.

[7] S. Liu, Y. Chen, W. Trappe, L. Greenstein, "ALDO: An Anomaly Detection Framework for Dynamic Spectrum Access Networks," in *IEEE INFOCOM 2009*, pp.675-683, 2009.

[8] N. R. Adam and J. C. Wortmann, "Security-control methods for statistical databases: A comparative study," *ACM Computing Surveys*, vol. 21, no. 4, pp. 515–556, 1989.

[9] T. Dalenius, "Finding a needle in a haystack - or identifying anonymous census records," *Jour. Off. Stats.*, vol. 2, no. 3, pp. 329–336, 1986.

[10] A. Dobra, S. Fienberg, and M. Trottini, *Assessing the Risk of Disclosure of Confidential Categorical Data*. Oxford University Press, 2000, vol. 7, pp. 125-144.

[11] C. Dwork, “Differential privacy,” in *Proc. 33rd Intl. Colloq. Automata, Lang., Prog.*, Venice, Italy, July 2006.

[12] L. Sankar, S. R. Rajagopalan, and H. V. Poor, “A theory of privacy and utility in databases,” under revision, *IEEE Trans. Inform. Forensics and Security, Special Issue on Privacy in Cloud Management Systems*.

[13] L. Sweeney. *k*-anonymity: A model for protecting privacy. *Intl. J. Uncertainty, Fuzziness, and Knowledge-based Systems*, 10(5):557–570, 2002.

[14] H. Yamamoto, ”A source coding problem for sources with additional outputs to keep secret from the receiver or wiretappers,” *IEEE Trans. Inform. Theory*, vol. 29, no. 6, pp. 918–923, Nov. 1983.

[15] C. Dwork, F. McSherry, K. Nissim, and A. Smith, ”Calibrating noise to sensitivity in private data analysis,” in *Proc. 3rd IACR Theory Crypto. Conf.*, New York, NY, Mar. 2006

[16] R. Agrawal, A. Evfimievski, and R. Srikant, “Data security and protection: Information sharing across private databases,” in *Proc. ACM Intl. Conf. Management Data*, 2003, pp. 86–97.

[17] R. C. W. Wong and E. Lo, “Competitive privacy: Secure analysis on integrated sequence data,” in *Proc. Database Syst. Ad. Apps.*, Tsukuba, Japan, Apr. 2010, pp. 168–175.

[18] L. Sankar, S. Kar, R. Tandon, and H. V. Poor, “Competitive privacy in the smart grid: An information-theoretic approach,” in *Proc. 2nd IEEE Intl. Conf. Smart Grid Commun.*, Brussels, Belgium, Oct. 2011.

[19] E. V. Belmega, L. Sankar, H. V. Poor, and M. Debbah, ”Pricing mechanisms for cooperative state estimation,” in *IEEE Intl. Symp. Commun., Control and Signal Proc. (ISCCSP)*, Roma, Italy, May 2012, pp. 1–4.

[20] L. Varshney and D. Oppenheim, ”On cross-enterprise collaboration,” in *Business Process Management*, vol. 6896,2011, pp. 29–37.

[21] E. V. Belmega, L. Sankar, and H. V. Poor, ”Repeated games for privacy-aware distributed state estimation in the smart grid,” in *Intl. Conf. on Network Games, Control and Optimization*, Avignon, Framce, Oct 2012.

[22] G. Kalogridis, C. Efthymiou, S. Z. Denic, T. A. Lewis, and R. Cepeda, “Privacy for smart meters: Towards undetectable appliance load signatures,” in *Proc. IEEE 1st Intl. Conf. Smart Grid Comm.*, Gaithersburg, MD, Oct. 2010, pp. 232–237.

[23] D. Varodayan and A. Khisti, “Smart meter privacy using a rechargeable battery: minimizing the rate of information leakage,” in *Proc. IEEE Intl. Conf. Acous., Speech, and Signal Proc.*, Prague, Czech Republic, 2011.

[24] V. Rastogi and S. Nath, “Differentially private aggregation of distributed time-series with transformation and encryption,” in *Proc. 2010 Intl. Conf. Data Management*, Indianapolis, Indiana, USA, 2010, pp. 735–746.

[25] L. Sankar, S. R. Rajagopalan, S. Mohajer, and H. V. Poor, “Smart meter privacy: A theoretical framework,” *IEEE Trans. Smart Grid*, no. 99, pp. 1 –10, 2012, early access article.

[26] L. Lai, S.-W. Ho and H. V. Poor, “Privacy-Security Tradeoffs in Biometric Security Systems - Part I: Single-Use Case,” *IEEE Trans. Information Forensics and Security*, Vol. 6, No. 1, pp. 122 - 139, March 2011.

[27] S. Reardon. FBI launches $1 billion face recognition project. NewScientist, Sept. 2012.

[28] Z. Erkin, M. Franz, J. Guajardo, S. Katzenbeisser, I. Lagendijk, and T. Toft, “Privacy-preserving face recognition,” in *LNCS*, vol. 5672, 2007, pp. 235–253.

[29] G. Fanti, M. Finiasz, and K. Ramchandran, “Private media search on public databases,” *Submitted*.

[30] C. Yeo, P. Ahammad, H. Zhang, and K. Ramchandran. ”Rate-efficient visual correspondences using random projections.” *In Proc. IEEE Intl. Conf. on Image Processing*, pp. 217–220, San Diego, CA, 2008.

Biographies

Lalitha Sankar is an Assistant Professor in Electrical, Computer, and Energy Engineering at Arizona State University. Her research interests include information privacy and secrecy in distributed and cyber-physical systems, wireless communications, and network information theory. Dr Sankar was a recipient of a three year Science and Technology Teaching Postdoctoral Fellowship from Princeton University. For her doctoral work, she received the 2007 Electrical Engineering Academic Achievement Award from Rutgers University. She received the IEEE Globecom 2011 Best Paper Award.

Mérouane Debbah is a Full Professor at Supelec (Gif-sur-Yvette, France), holder of the Alcatel-Lucent Chair on Flexible Radio and a recipient of the ERC starting grant MORE. His research interests are in information theory, signal processing and wireless communications. He is an Associate Editor for IEEE Transactions on Signal Processing. Mérouane Debbah is the recipient of the ”Mario Boella” award in 2005, the 2007 General Symposium IEEE GLOBECOM best paper award, the Wi-Opt 2009 best paper award, the 2010 Newcom++ best paper award as well as the Valuetools 2007, Valuetools 2008 and CrownCom2009 best student paper awards. In 2011, he received the IEEE Glavieux Prize Award.

H. Vincent Poor is the Michael Henry Strater University Professor of Electrical Engineering at Princeton University, where he is Dean of the School of Engineering and Applied Science. His research interests include information theoretic privacy and security in wireless networking and related fields. He is a member of the National Academy of Engineering, the National Academy of Sciences, and is a Fellow of the IEEE, the American Academy of Arts and Sciences, the Royal Academy of Engineering (U.K.) and other scientific and technical organizations. Dr. Poor is currently serving on the Editorial Board of the IEEE Signal Processing Magazine. He received the Technical Achievement and Society Awards of IEEE Signal Processing Society in 2007 and 2011, respectively, as well as the 2010 IET Ambrose Fleming Medal and the 2011 IEEE Eric E. Sumner Award.

Kannan Ramchandran is a Professor of Electrical Engineering and Computer Science at the University of California at Berkeley. He is a Fellow of the IEEE, and has won numerous research awards including several Best Paper awards from the IEEE Signal Processing Society, and the IEEE Communications and Information Theory Societies, an Okawa Foundation Research Prize at Berkeley, a Hank Magnusky Scholar award at the University of Illinois, and an Eli Jury thesis award at Columbia University. He was also recognized with a Distinguished Teaching Award from his department at UC Berkeley. His research interests include distributed and adaptive signal processing for large-scale systems, codes for distributed storage systems, peer-to-peer networking and multimedia content delivery, and multi-user information and communication theory, with an emphasis on security and privacy.

Wade Trappe is an Associate Professor in the Electrical and Computer Engineering Department at Rutgers University, and Associate Director of the Wireless Information Network Laboratory (WINLAB), where he directs WINLAB’s research in wireless security. Professor Trappe has served as an editor for IEEE Transactions on Information Forensics and Security (TIFS), IEEE Signal Processing Magazine (SPM), and IEEE Transactions on Mobile Computing (TMC). He served as the lead guest editor for September 2011 special issue of the Transactions on Information Forensics and Security on “Using the Physical Layer for Securing the Next Generation of Communication Systems” and also served IEEE Signal Processing Society as the SPS representative to the governing board of IEEE TMC.