This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Wavelet-Based Multiresolution Smart Meter Privacy

Dominik Engel, *Member, IEEE*, and Günther Eibl, *Member, IEEE*

***Abstract*—The availability of individual load curves per household in the smart grid end-user domain combined with non-intrusive load monitoring (NILM) to infer personal data from these load curves has led to privacy concerns. Based on insights of the interrelation of load profile resolution and accuracy of NILM techniques, we propose the use of the wavelet transform to represent load data in multiple resolutions. Each resolution is encrypted with a different key using an appropriate cipher and a hierarchical keying scheme. End-to-end security ensures access control. To meet requirements of low computational complexity in low-cost smart meters, the lifting implementation of the wavelet transform is used to generate multiple resolutions. It is shown that the multiresolution approach is compatible with other privacy-enhancing technologies, such as secure signal processing. This allows adding new degrees of freedom to these methods by introducing the dimension of multiple resolutions. The proposed approach is evaluated based on the provided level of privacy and security, computational demands, and feasibility in an economic sense.**

***Index Terms*—Privacy, smart metering, wavelet transform, multiresolution, conditional access.**

I. Introduction

SMART METERS form a central component of the smart grid and in combination with consumer energy management systems (CEMS) provide an interface to smart home technology. Each smart meter is capable of measuring, storing and transmitting detailed load profiles. Typically, the data is transmitted on a daily basis. The exact granularity of the transmitted load profiles is not finally specified, and may differ by country. The intervals between single measurements will lie between a few seconds and several minutes.

The deployment of smart meter technology and the ensuing availability of fine-grained consumption data has led to severe privacy concerns. Already in the 1990s, Hart [1] showed that personally sensitive information can be extracted from load profiles through so-called "non-intrusive load monitoring" (NILM). More recent studies improved on the methods, e.g., Lisovich *et al.* [2] showed that the appliance information could be used to infer personal information, such as sleep-wake-cycles and presence, but in theory also lifestyle and religion.

Manuscript received April 2, 2015; revised July 13, 2015 and October 12, 2015; accepted November 22, 2015. This work was supported in part by the Austrian Federal Ministry of Science, Research, and Economy; in part by the Austrian National Foundation for Research, Technology, and Development; and in part by the Federal State of Salzburg. Paper no. TSG-00376-2015.

The authors are with the Josef Ressel Center for User-Centric Smart Grid Privacy, Security, and Control, Salzburg University of Applied Sciences, Puch/Salzburg A-5412, Austria (e-mail: dominik.engel@en-trust.at).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSG.2015.2504395

The fact that accuracy of detection heavily depends on the resolution of the investigated load profile is often neglected. Consider the results by Greveler *et al.* [3], who found that for some TV sets the multimedia content could be determined by smart meter data at a resolution of 2 seconds. These results were incorrectly generalized by mainstream media and scientific contributions alike without regard for resolution.

In [4], Eibl and Engel report results of a first systematic investigation of the influence of resolution on smart meter privacy. It is shown that the intuitive expectation that the accuracy of NILM methods decreases with resolution can also be motivated systematically. It is shown that decreasing the resolution of load profiles transmitted by a smart meter increases privacy. It is clear that the requirements of smart grid use cases with respect to resolution differ greatly (e.g., billing only requires a very low resolution, network monitoring requires a higher resolution and using NILM methods for energy disaggregation to provide energy saving advice will require an even higher resolution). Furthermore, it is clear that putting control over which data resolution to send to which stakeholder into the hands of the end user will dramatically increase user acceptance.

In this paper, we propose a system for privacy-preserving smart metering. It gives end-users control over access to their load profiles in different resolutions. Thereby, a user-centric privacy approach is realized. Furthermore, limitation of resolution can be done in the encrypted domain. The system integrated previously presented methods [5]–[7] for smart meter privacy based on the wavelet transform into a comprehensive framework, which takes the recent results on the impact of resolution on privacy into account [4]. We discuss how the pieces can be put together, and what privacy use cases can be realized with the integrated approach. The following requirements are met by this system:

- Multi-resolution representation without data expansion,
- Low computational overhead,
- Conditional access to each resolution,
- Preservation of sum over all resolutions (to support, e.g., billing),
- Compatibility with other Privacy-Enhancing-Technologies (PETs), such as secure homomorphic aggregation, per resolution to add additional privacy choices, and
- Compatibility with hierarchical key generation.

A main contribution is the transfer and adaptation of methods from other problem domains to the area of smart metering to create a comprehensive smart metering approach that can balance both, requirements for functionality and privacy. The choice and combination of methods is one aspect of this contribution, the adaptation and tailoring of the individual methods use in smart metering is another aspect. Finally, to evaluate

1949-3053 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

the effectiveness of the approach, a privacy measure for smart metering is introduced.

The remainder of this paper is organized as follows: Section II gives an overview of related work and discusses suggestions for other privacy-enhancing technologies. The impact of resolution on privacy is reviewed in Section III-A. Section III contains all details on the proposed wavelet-based approach for smart meter privacy. This section also shows that the used wavelet transform preserves the sum over all resolutions, which is an important property for use cases like billing. The proposed approach is evaluated in Section IV. In Section V the compatibility with other PETs is discussed. Section VI concludes and gives an outlook on future work.

II. Related Work

There are a number of contributions for privacy-enhancing technologies for smart metering. Jawurek *et al.* [8] argue, that approaches relying on policy alone, may prove inadequate to provide a sufficient level of privacy and that technological methods that enforce privacy by virtue of "strength of mechanism" need to be employed. Indeed, a number of such technological approaches have been suggested to remedy the loss in privacy and still enable smart metering functionality on a broad basis. In the following, we give a brief overview of these contributions, based on [9]. More detailed surveys can be found in [8], [10], and [11].

The only approach that is widely used in the real world at this point in time, is *anonymization or pseudonymization* of smart metering data. Consumption data and the personal data are split and stored separately. Methods for de-anonymization are a major threat for these types of approaches. It has been shown that even after anonymization or pseudonymization, data items can still be attributed to the individual that originated them. Jawurek *et al.* [12] show that de-anonymization can also be done in the smart grid user domain. This structural traceability is a problem for schemes that rely on anonymization or pseudonymization only without the use of additional encryption.

Simple aggregation tries to hide data related to individuals by aggregating over a number of house-holds, e.g., all households in a neighborhood are network (NAN). For example, Bohli *et al.* [13] propose a privacy scheme in which high resolution smart meter readings are aggregated at NAN level and only the aggregate is sent to the utility. They introduce two solutions both with and with-out involvement of trusted third parties.

Due to the inherent link between load data resolution and privacy, splitting the load data into a variety of *different resolutions*, each associated with different authorization levels, has been proposed by a number of contributions. For example, the anonymization scheme proposed by Efthymiou and Kalogridis [14] is based on two different resolutions: a low resolution that can be used for billing purposes, and a high resolution that allows further investigation. This scheme employs a trusted third party escrow service. In the manuscript presented here, we build on previous work on wavelet-based multi-resolution privacy [6], [7].

Masking relates to approaches which add numerical artifacts, e.g., random sequences to the original load data to obfuscate individual contributions. The added artifacts are constructed in such a way that they cancel each other out upon aggregation. The aggregator can therefore combine the data values of all participant to create an accurate aggregation, but cannot gain access to individual contribution. For example, Kursawe *et al.* [15] propose such an aggregation protocol, which compared to other approaches has the advantage of relatively low computational complexity. Defend and Kursawe [16] further improve on this idea. Danezis *et al.* [17] present another low-overhead protocol for aggregation of smart meter data, which puts minimal computational demands on the smart meter hardware.

Differential privacy, as Dwork [18, p. 1] puts it, roughly speaking, "ensures that (almost, and quantifiably) no risk is incurred by joining a statistical database". Adding or removing an item from the database will not (or only to a very limited degree) affect the result of statistical computations. This is commonly achieved by the distributed generation of noise which is added to the individual data contribution. Shi *et al.* [19] propose a scheme for adding random noise to time series data using a symmetric geometric distribution. An advantage of this scheme is that the participants need not trust each other, nor rely on a trusted aggregator. As another example, Ács and Castelluccia [20] obscure individual data sets by adding Laplacian noise, which is jointly generated by the participants. Apart from the obvious drawback that the data is no longer exact after differential privacy is applied, data pollution by malicious participants is another issue with this approach [19].

Secure Signal Processing (SSP) refers to the possibility to perform certain computations, such as aggregation in the encrypted domain. A commonly employed mechanism in SSP is homomorphic encryption, which allows some specific manipulations of the ciphertext to be reflected in the plaintext domain. For example, Li *et al.* [21] propose an overlay network in a tree-like topology and the use of a Paillier cryptosystem. Garcia and Jacobs [22] combine secret sharing with a Paillier cryptosystem to add flexibility in the aggregation (at the expense of additional computational complexity). Erkin and Tsudik [23] extend the idea of homomorphic encryption of smart meter readings by splitting the module into random shares, which, in combination with a modified Pailler cryptosystem, allows flexible spatial and temporal aggregation for different use cases, such as billing or network monitoring.

III. Wavelet-Based Smart Meter Privacy

A. Motivation for Multi-Resolution Privacy

The basis for both, regulatory-based and technology-based approaches, is detailed knowledge of what information can be extracted from the available user data. To date, there is little systematic research on this subject in the context of smart grids.

In [24], Molina-Markham *et al.* investigate the information revealed from load profiles at different granularities.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

They show that with off-the shelf statistical methods detailed information on the behavior of users can be inferred from load profiles without prior knowledge or precomputed appliance signatures. They argue that "the information leaks directly correlate with the time granularity that a meter measures power consumption" [24, p. 61] and list a number of privacy-relevant questions that can be answered using load profiles at granularities ranging from hours to seconds.

In [4], Eibl and Engel study the impact of resolution on NILM methods systematically. The authors use the so-called F-Score, a combination for accuracy and precision to evaluate appliance detection in a number of different resolution. The approach is empirically evaluated using the publicly available REDD energy disaggregation data set [25]. The results clearly and systematically show that appliance detection accuracy decreases with resolution.

Based on these insights, it is only logical to aim at differentiating access control to load profile by resolution. The classical wavelet transformation in the lifting implementation is the ideal tool to create integrated, multi-resolution representations of load profiles. In contrast to the Fourier transform, which only provides localization in the frequency domain, the wavelet transform allows to strike a balance between localization in the time domain and the frequency domain. This allows to create a cascade of lower resolution representations of the original sequence, each of which exactly correspond to a sub-sampled version of the previous resolution with half its size. Each resolution contained in the multi-resolution load profile can be tailored to correspond to a class of detection accuracy (as will be discussed in Section IV-B). Granting access to third party based on this multi-resolution representation allows informed, privacy-aware data exchange to the user.

Multi-resolution privacy, as proposed here can be seen as orthogonal to many of the approaches mentioned in Section II, i.e., approaches such as secure signal processing or masking can be combined with and enhanced by multi-resolution analysis. This leads to more possibilities that can be presented to the user, e.g., a higher resolution could be restricted to be communicated over secure aggregation only, while a lower resolution could be communicated as an individual load profile to a defined external party through end-to-end security. Note that the approach proposed here supports both, billing and aggregation use-cases (which are often distinguished in literature). The combination of multi-resolution privacy with selected other PETs is discussed in more detail in Section V.

B. Multi-Resolution Load Profile Representation

Let $L[i]$ be a tuple of length n, with $i = 1, \ldots, n$, containing the data values of the original load profile (without loss of generality, we assume these data values to be of type float). For the sake of readability, we omit the index and write $\mathbf{L}$ whenever we do not need to refer to individual elements of $L[i]$.

A wavelet transform of maximum depth d, denoted as $W_d(\mathbf{L})$, is applied to the original load profile $\mathbf{L}$ by iteratively applying the transformation in d steps to the resulting low pass subbands: In each step k, for $k = 1, \ldots, d$, the wavelet transform operates on data of resolution $r \coloneqq d - k$ and produces

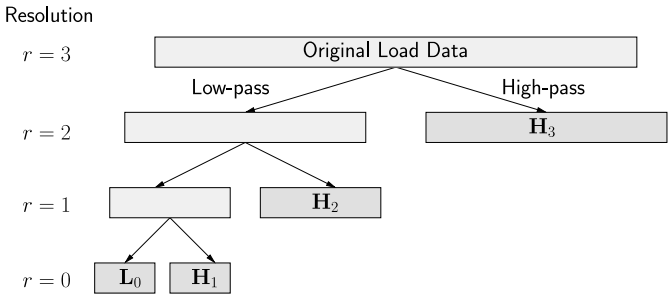


Fig. 1. Wavelet transformation of a smart meter load profile.

the next lower resolution, which is half the size of the data at resolution r. As illustrated in Fig. 1, in each step k, a low-pass subband $\mathbf{L}_{r-1}$ and a high-pass subband $\mathbf{H}_r$ are produced in this way, each half of the size of the input data, i.e., the number of samples after step k is $n \cdot 2^{-k}$. The coefficients contained in $\mathbf{H}_r$ are stored as the wavelet coefficients of resolution r. Then the wavelet transform is applied to $\mathbf{L}_{r-1}$ to produce the next lower resolution, until the lowest resolution $r = 0$ is reached. At the end of the transformation the sequence $\mathbf{L}_0,\ \mathbf{H}_1, \ldots,\ \mathbf{H}_d$ is obtained, see Fig. 1. In the following, we use $\mathbf{w}_i$ as a short hand to refer to each of the individual subbands $\mathbf{L}_0, \mathbf{H}_1, \ldots, \mathbf{H}_r\}$, i.e., $\mathbf{w}_0 \coloneqq \mathbf{L}_0$, $\mathbf{w}_1 \coloneqq \mathbf{H}_1, \ldots, \mathbf{w}_d \coloneqq \mathbf{H}_d$. Furthermore, we denote the *all* wavelet coefficients necessary for a resolution r by $\mathbf{W}_r$, i.e., $\mathbf{W}_r \coloneqq \{\mathbf{L}_0, \mathbf{H}_1, \ldots, \mathbf{H}_r\}$. For a wavelet transform of depth d, the maximum (i.e., original) resolution is d.

The synthesis step of the inverse wavelet transform W_d^{-1} starts with the lowest resolution $r = 0$. To get the next higher resolution of the signal the next higher resolution subband is needed, so that in a series of d steps one finally obtains the original load curve (since we only consider lossless transformations).

In order to provide a load profile $\mathbf{L}_t$ with a target resolution r from coefficients $\mathbf{W}_d$ (where $r \leq d$) only r synthesis steps need to be performed and only the subbands with resolution $t \leq r$, i.e., $\mathbf{W}_r$ are needed. Denoting the selection of the lower r resolutions from $\mathbf{W}_d$ as $T_r(\cdot)$ this can be written as

$$\mathbf{L}_r = W_r^{-1}(T_r(\mathbf{W}_d)) \tag{1}$$

The operator T_r can be generalized to be any linear transformation T of the wavelet coefficients to be used for example for denoising, which could be valuable for transmission of signal aggregations. Representing load profiles in the wavelet domain can also be a basis for signal processing [26] and data compression [27], [28]. As these aspects are out of the scope of this paper, the combination of signal processing and compression with the privacy approach proposed here will be addressed by future work. Another item to be considered for future work is the use of non-uniform sampling: for areas where the signal is smooth the sampling interval could be increased adaptively, whereas for areas where the signal is less smooth, the sampling interval could be increased – in this way privacy could be increased while still retaining the same average sampling rate.

The discrete wavelet transform does not lead to data expansion, and by using a lifting implementation the transformation can even be done in-place: "The lifting scheme

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

also leads to a fast in-place calculation of the wavelet transform, i.e., an implementation that does not require auxiliary memory." [29, p. 4]. Therefore, the number of bits needed to represent the coefficient data of a level d wavelet transform, $\mathbf{W}_d$, is the same number of bits needed to represent the original load data L.

The wavelet coefficients of the different subbands can be represented in a single, embedded bitstream, which correspondingly contains all resolutions. Note that if the appropriate filter is used, the wavelet transform is lossless, i.e., no data loss occurs and the original load curve can be recovered perfectly from the coefficients contained in the embedded bitstream.

To implement multi-resolution analysis in a manner that is suitable for smart metering devices, wavelet lifting [30] provides a helpful perspective: This view on the wavelet transform factors wavelet filters into lifting steps, which for many filters rely on simple operations only.

C. Applying the Haar Wavelet to Load Profiles

In this paper, we use the simple Haar wavelet filter to create multi-resolution load profiles. Other filters have been studied by Engel in [6], where the author came to the conclusion that the Haar wavelet is sufficient for all currently envisioned use cases and at the same time has the important property of very low computational complexity: The Haar wavelet filter realizes low-pass filtering as averaging of the sample values. The high-pass step is realized by calculating the corresponding differences to allow for lossless reconstruction.

Let $\mathbf{L}_r$ be the input signal at resolution r, and $\mathbf{L}_{r-1}$ and $\mathbf{H}_r$ be the low-pass and high-pass output signals, respectively. Further let n be the length of $\mathbf{L}_r$, and for the sake of simplicity, let $n \bmod 2 = 0$. The lifting steps for the forward transform (going from resolution r to the next lower resolution $r-1$) with the Haar wavelet can be written as follows (adapted from [30]):

$$\hat{L}_{r-1}[i] = L_r[2i] \tag{2}$$

$$\hat{H}_r[i] = L_r[2i+1] \tag{3}$$

$$H_r[i] = \hat{H}_r[i] - \hat{L}_{r-1}[i] \tag{4}$$

$$L_{r-1}[i] = \hat{L}_{r-1}[i] + \frac{1}{2}(H_r[i]), \tag{5}$$

with $i = 1, \ldots, \frac{n}{2}$. The inverse transform (going from resolution $r-1$ to $r-1$) correspondingly is given as

$$\hat{L}_{r-1}[i] = L_{r-1}[i] - \frac{1}{2}(H_r[i]) \tag{6}$$

$$\hat{H}_r[i] = H_r[i] + \hat{L}_{r-1}[i] \tag{7}$$

$$L_r[2i+1] = \hat{H}_r[i] \tag{8}$$

$$L_r[2i] = \hat{L}_{r-1}[i] \tag{9}$$

again with $i = 1, \ldots, \frac{n}{2}$.

This transformation is lossless, and the applied operations in each step are equivalent to subsampling. In effect, each iteration of applying the Haar wavelet is equivalent to halving the sampling rate.

The Haar wavelet filter perfectly preserves the first moment in each step of the transform. The sum of the original load profile (i.e., the total consumption) can be accessed at any resolution, because, as can be easily shown, the sum for the load profile at any resolution r can be obtained from the next lower resolution $r-1$ as follows:

$$\sum_{i=1}^{n} L_r[i] = 2 \cdot \sum_{j=1}^{n/2} L_{r-1}[j]. \tag{10}$$

This is an important property, as it allows the use of any lower resolution for functions like accurate billing, as the sum of the original sequence can be derived from any of the lower resolutions.

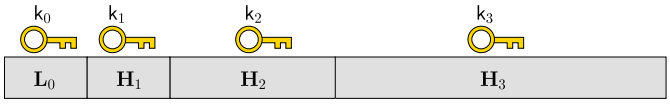


Fig. 2. Final encrypted bitstream produced by the smart meter: The wavelet coefficients of each subband are encrypted with an individual key.

D. User-Centric Privacy Through Conditional Access

The idea of conditional access stems from the context of multimedia entertainment data. Entertainment content usually exists in various resolutions (e.g., mobile content, standard definition, high definition), which may be priced differently. A multi-resolution representation of the multimedia data allows the efficient representation of the resolutions in a single bitstream. This is an advantage as only one version of the bitstream needs to be handled and transmitted. Conditional access allows users to pay only for the resolutions they are interesting in. For example, the owner of a standard definition television has no need to pay for the high-definition version of the content. Through conditional access only the bitstream portion relevant for the desired resolution is decrypted, the rest of the bitstream is ignored.

We propose to use the conditional access paradigm for smart metering data in multi-resolution representation. Each subband $\mathbf{L}_0, \mathbf{H}_1, \ldots, \mathbf{H}_d$ is encrypted with a different key (key generation and handling are discussed in Section III-E). The whole datastream is transmitted over a Smart Grid communication infrastructure. Access to the different resolutions is thereby only granted to parties that hold the needed keys, as illustrated by Figure 2.

This scheme allows flexible control by the end-user how access is granted to smart meter data. For example, a particular energy provider may be granted access only to the lowest resolution for billing purposes. A third-party service providing energy saving advice by employing NILM methods may be granted access to the highest resolution by the user. The end-user may further be willing to provide data for network monitoring, but only at a medium resolution. Note that the approach presented here provides all the necessary means and ingredients for multi-resolution privacy, but it does not make the decision on which resolution to choose on the users' behalf. This part could be provided, e.g., by a recommender system, which advises the user on the privacy implications of a certain resolution. First steps into such automated privacy recommendations have been made in [31].

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

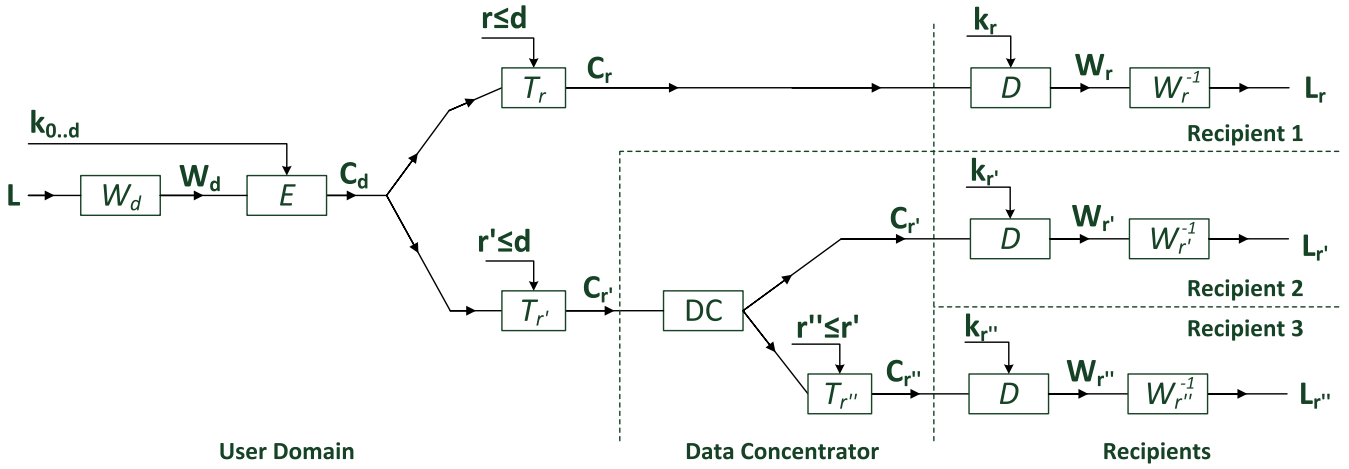


Fig. 3. Illustration of Wavelet-based Multi-Resolution Privacy: Three communication paths from user domain to recipient.

T_t (where $0 \leq t \leq R$), i.e., decreasing the resolution, can be applied before or after encryption. Note that no keys are required to decrease the resolution of the encrypted bitstream, as the operation is a simple truncation: For reducing the resolution from r to $r-1$, discarding the bits of the encrypted coefficients of $\mathbf{H}_r$ is sufficient. As each resolution is encrypted separately, the boundaries are known and the truncation can be done in the encrypted domain (if a cipher is used that preserves the number of bits, the truncation points are known implicitly, otherwise explicit markers can be introduced).

This allows rate adaptation not only by the users themselves, but also at a later point, e.g., through a third party (no keys need to be provided for this third party). In this way, the proposed scheme also enables relaying of data in various resolution. An illustrative example is discussed in Section III-F.

E. Hierarchical Key Generation

In [32], a scheme for generating keys for multi-resolution privacy is proposed, based on previous suggestions stemming from multimedia security (e.g., [33]). The original idea is due to Lamport [34], who already proposed the underlying method in 1981.

For each resolution r, the key k_{r-1} for the next lower resolution $r-1$ is obtained by using a cryptographic one-way hash function h on k_r:

$$\mathsf{k}_{r-1} = h(\mathsf{k}_r). \tag{11}$$

In this way, for each resolution r, all lower resolution keys $\mathsf{k}_{r-1}, \mathsf{k}_{r-2}, \ldots, \mathsf{k}_0$ can always be obtained. Using this hierarchical key generation scheme saves overhead in key management, as only a single key needs to be stored and transmitted.

By using a secure cryptographic one-way hash function, the one-way property ensures that inferring keys for higher resolutions from a lower resolution key is extremely difficult.

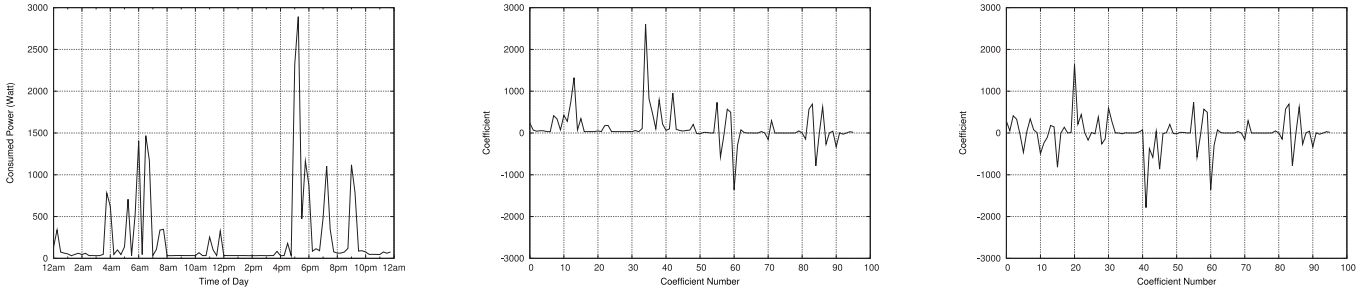
Note that the hierarchical keys are generated for a symmetric scheme, such as AES. Access to the different resolution to different stakeholders is granted by using the public keys of these stakeholders. For examples, if a user wants to grant access to an external party to resolution $r = 2$, the user encrypts k_2 with the public key pk of this external party, producing a "wrapped" key $\mathsf{wk} = E_{\mathsf{pk}}(\mathsf{k}_2)$. The external party can use its private key sk to obtain the symmetric key: $D_{\mathsf{sk}}(\mathsf{wk}) = \mathsf{k}_2$. Subsequently, keys k_1 and k_0 can be derived by using Equation 11. With these keys, all wavelet coefficients of $\mathbf{L}_0, \mathbf{H}_1$ and $\mathbf{H}_2$ can be decrypted. Finally, the inverse wavelet transform is applied to obtain the load profile in the target resolution.

F. Illustration

Figure 3 illustrates the proposed method. In the user domain, a wavelet transform of depth d is applied to the original load profile $\mathbf{L}$, resulting in the wavelet coefficients $\mathbf{W}_d$. The coefficients for each resolution $r = 0, \ldots, d$ are encrypted with a unique key k_r (using the hierarchical key scheme, as discussed above). Access to different resolutions is granted to recipients based on these keys. In the illustration, there are three recipients, each of which is granted access to a different resolution, r, r', and r'', respectively. Access to resolution r, the highest resolution, is granted to Recipient 1, which could be a third party service provider for energy optimization through NILM. Resolutions r' and r'' are relayed over a data concentrator (typically operated by the DSO). Recipient 2 could be the DSO itself, which is granted access to resolution r' to use the data at this resolution for demand prognosis. Recipient 3 could be the energy provider, which receives a very low resolution r'' through the DSO's data concentrator for billing purposes. (Note that the use of a data concentrator is possible with the proposed scheme, but not required. The regulation in some countries prescribes the role of a data concentrator, in other countries no data concentrators are used.)

Before the data leaves the user domain, by applying T the user can decrease the data to be transmitted. In the illustration, there are two communication channels which leave the user domain: the upper communication channel is a *direct channel* to Recipient 1 (e.g., via the user's Internet connection). For this recipient, the user grants access to resolution r by providing key k_r. Furthermore, only the coefficient data up to resolution r needs to be transmitted. This can be achieved

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



(a) Original Data (b) Level 1 Decomposition (c) Level 5 Decomposition

Fig. 4. Example for Wavelet Decomposition of Smart Meter Data.

by creating $\mathbf{C}_r$ through applying T_r to $\mathbf{C}_d$. Recipient 1 can decrypt $\mathbf{W}_r$ from C_r and by applying the inverse wavelet transform can reconstruct $\mathbf{L}_r$.

The lower communication channel leaving the user domain transmits encrypted wavelet coefficient data to Recipient 2 and Recipient 3, via a *data concentrator* (DC). As the maximum resolution for all recipients behind the DC is r', the transmitted data can be reduced to $C_{r'}$ before it is passed on to the DC.

The DC does not have any keys. Nevertheless, the DC can apply T to $\mathbf{C}_{r'}$ to reduce the amount of data transmitted to Recipient 3 to the lower resolution r''. The option of nodes in the network being able to perform rate adaption is advantageous in networks with low bandwidth links, such as narrow-band powerline communication (PLC). Note that the reduction of resolution by the data concentrator DC by applying $T_{r''}$ is not relevant for security. Even if DC had passed on $\mathbf{C}_{r'}$ to Recipient 1 instead of $\mathbf{C}_{r''}$, Recipient 1 would still lack the key to decrypt $\mathbf{W}_{r'}$.

Both, Recipient 2 and Recipient 3 can obtain the load profiles $\mathbf{L}_{r'}$ and $\mathbf{L}_{r''}$, respectively, in the resolutions intended by the user.

To also illustrate the data perspective, Figure 4 shows an example of the wavelet decomposition of actual data. The original sampling interval in this example is 15 minutes, i.e., 96 values per day, as shown in Fig. 4(a). The coefficients of a level-1 decomposition with the Haar wavelet is shown in Fig. 4(b). The left half of these coefficients represent an lower resolution of the original sequence with a sampling interval of 30 minutes. Fig. 4(c) shows a level-5 wavelet decomposition. From this representation, 6 different resolution (including the original resolution) can be obtained, with the lowest resolution being composed of the three left-most values and representing a sampling interval of 8 hours.

IV. Evaluation

In this section, the proposed scheme is evaluated with respect to computational demands, security and privacy and real-world feasibility.

A. Complexity

As discussed above, implementing the wavelet transform as lifting steps is computationally inexpensive. Generally, the discrete wavelet transform has a complexity of $\mathcal{O}(n)$. Due to the simple operations used in the lifting implementation, the transformation part can be realized by inexpensive smart meter hardware.

The computational demands for encryption depends on the used encryption scheme. For standard encryption schemes, efficient implementations exist that can be integrated into smart meter hardware. Some overhead is introduced for key management, and potentially for the creation of session keys.

The proposed method has been implemented as a proof of concept in Java (Oracle Java v8 with ARM-extensions, version 1.8.0 with hard float). The method was evaluated in a low-cost ARM-based environment (Raspberry Pi 2, featuring a 900MHz quad-core ARM Cortex-A7 CPU and 1 GB RAM at a cost of $35) running Raspbian Linux (in the version released on February 15, 2015, based on Debian Wheezy). The choice of this hardware platform is sensible, as it reflects the computing capabilities smart metering hardware will most likely provide. Rather than choosing the current solution of a single smart meter manufacturer, we use the open Raspberry Pi platform in combination with Linux to provide a test environment that is representative of future smart meters in a more general way.

To evaluate the method, we use the publicly available REDD data set [25]. This data set contains load data for a number of houses over the period of several weeks at a measuring interval of 3 seconds. For our test, we use 14 days of load profiles from house 1. We use the first 28,672 samples of the data set for each day, which corresponds to a measuring interval of 3.01 seconds. This sampling interval allows us a maximum wavelet decomposition depth of 12 without the need for border handling (because the number of samples equals $7 \cdot 2^{12}$), with the lowest resolution having a size of 7 samples (i.e., one aggregated value for every 3.4 hours). The sizes of the resolutions are given in Table I. Note that in real world setups, the number of measurements per day can be chosen by the smart meter, but will be affected by local legislation.

The following encryption scenarios were used: (1) wavelet transform only without any encryption, (2) Symmetric encryption: AES with 128-bit and 256-bit keys, and (3) Hybrid encryption: 128-bit and 256-bit AES resolution keys encrypted with 2048-bit RSA keys.

The following steps are executed in the scenarios: (i) Apply a Haar wavelet transform of depth 12 to the load profile (all scenarios), (ii) Generate 13 hierarchical AES resolution

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
RESOLUTION SIZES AND TEST RESULTS; X: ACTIVITY CAN BE INFERRED FROM DATA, C: COOKING, B: BATHROOM ACTIVITIES, H: HOUSEWORK, P: PRESENCE/ABSENCE

Resolution	Size	Approx. Interval	C	B	H	P
12	28672	3s	X	X	X	X
11	14336	6s	X	X	X	X
10	7168	12s	X	X	X	X
9	3584	24s	X	X	X	X
8	1792	48s		X	X	X
7	896	1.6m				X
6	448	3.2m				X
5	224	6.4m				X
4	112	12.9m				X
3	56	25.7m				X
2	28	51.4m				X
1	14	1.7h				
0	7	3.4h				

TABLE II
EXECUTION TIME FOR MULTI-RESOLUTION ENCRYPTION ON RASPBERRY PI 2, AVERAGE FOR 14 LOAD PROFILES WITH 500 EXECUTIONS EACH

Scenario	Avg. Exec. Time (ms)	Std. Deviation
(1) WAV only	9.1	1.2
(2) AES 128/256	46.3 / 52.6	2.3 / 0.6
(3) HYB 128-2048/256-2048	173.4 / 179.9	1.8 / 1.5
Key generation AES 128/256	1388 / 1390	17

keys (Scenarios 2 and 3 only), (iii) Encrypt $\mathbf{L}_0, \mathbf{H}_1, \ldots, \mathbf{H}_{12}$, each with a different key (Scenario 2 and 3 only), and (iv) Encrypt the 13 resolution keys with a 2048-bit RSA public key (Scenario 3 only).

The results are shown in Table II. The timing results are given in milliseconds comparing wavelet transform only (WAV) with AES and hybrid encryption (HYB) using an AES session key encrypted with an RSA public key. In each category, the 14 daily load profiles from the REDD data set were investigated, each of which was transformed and encrypted 500 times. The results present the average time needed for processing one load profile (i.e., one day).

It can be seen that compared to the computational demands of the encryption stage, the computational demands for the wavelet transform are almost negligible. On average, the transformation of a load profile takes 9 ms. This fact is a strong argument in favor of the proposed approach. Considering other tasks smart meters will need to be able to handle (such as key management), multi-resolution support comes at practically no additional cost.

Some overhead is incurred by the need to create the resolution keys. For creating 13 session keys with the Java standard pseudo-random number generator (SHA1PRNG), on a Raspberry Pi 2 and averaged over 500 executions our implementation took approx. 1390ms.

B. Privacy and Security Analysis

In the following, we first review the proposed protocol in more formal detail, to make clear which party generates which keys and which party performs the encryption. We then outline the basic assumptions regarding adversary behavior. The important aspect of information reduction through subsampling is discussed. Finally, based on this discussion, the used notion of privacy is defined in more detail.

1) Protocol Review: The proposed protocol is given in Figure 5. For the discussion, one exemplary use case of the proposed method was selected: A smart meter collects energy consumption data, performs multi-resolution analysis followed by encryption and sends the ciphertext to a DSO. We also include the possibility of an optional concentrator, which can adapt the resolution by applying $T(\cdot)$ to the ciphertext.

The DSO's public key $\mathsf{pk}_{\mathrm{DSO}}$ is made available to the smart meter via a public key infrastructure (assuming that the smart meter is initially provisioned with the public keys of the used certifying authorities). Note that, for the sake of simplicity, basic security measures (such as authentication and integrity checking) are not discussed here, but of course should be added in a real-world application.

A wavelet transformation is performed by the smart meter resulting in d resolutions (line 1). The individual resolution keys k_i are created by the smart meter with the hierarchical keying scheme discussed in Section III-E (line 2). The coefficients of each resolution are encrypted with the corresponding resolution key (line 3), using a symmetric cipher (in our tests we use AES).

The smart meter is configured to grant the DSO access to the consumption data up to resolution r (with $r \leq d$). The corresponding resolution key k_r is therefore encrypted with the DSO's public key (line 5), producing the "wrapped" key $\mathsf{wk}_{\mathrm{DSO}}$. Note that due to using a hierarchical keying scheme only this single key k_r needs to be transmitted to the DSO (the necessary keys for the lower resolution coefficients can be derived from k_r). The encrypted coefficients $\mathbf{C_d}$, together with the wrapped resolution key $\mathsf{wk}_{\mathrm{DSO}}$, are transmitted by the smart meter. Note that, additionally, the smart meter could add more wrapped resolution keys, encrypted for other recipients with the corresponding public keys.

In the topology, an optional data concentrator can be used: The data concentrator receives ciphertexts by various smart meters and passes them on to the DSO. The data concentrator can be configured to perform resolution adaption. In the protocol in Figure 5, the data concentrator can truncate the resolutions higher than r by simply discarding the corresponding encrypted coefficients (as discussed above, no decryption is necessary here).

The DSO decrypts the resolution key k_r from $\mathsf{wk}_{\mathrm{DSO}}$ with its private key $\mathsf{sk}_{\mathrm{DSO}}$ (line 6) and derives the resolution keys for the lower resolutions (line 7). It can then decrypt the encrypted coefficients for the resolutions up to r (lines 8-9). Finally, by applying the inverse wavelet transform, the load data $\mathbf{L}_r$ of resolution r is obtained (line 10).

2) Adversary Model Assumptions: The smart meter is assumed to be trusted: It will reliably realize wavelet transform, key generation and encryption. It is assumed that no active or passive adversary has access to the internal processing of the smart meter.

The DSO and potential other legitimate recipients of the load data (in a specific resolution) are assigned an honest-but-curios (semi-honest) role: They will reliably perform wavelet

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

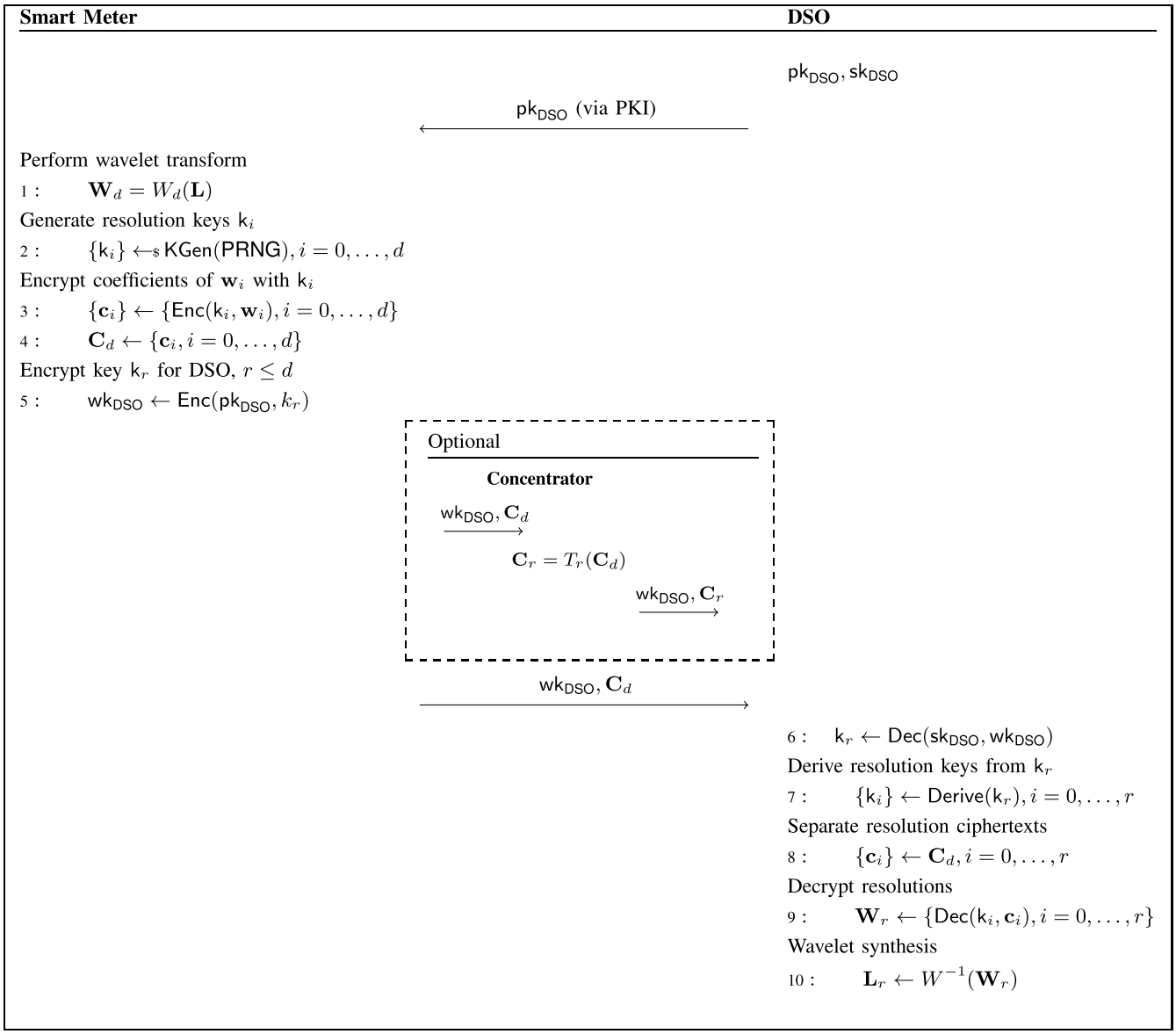


Fig. 5. Protocol for Multi-resolution Privacy for transferring a load profile $\mathbf{L}$ from the smart meter to the DSO at reduced resolution r as $\mathbf{L}_r$, with optional additional rate adaption applied by a data concentrator. $\mathsf{pk}_{\mathsf{DSO}}$ denotes the DSO's public key, $\mathsf{sk}_{\mathsf{DSO}}$ denotes the DSO's private key. k_i denotes the resolution keys used for symmetric encryption of the wavelet coefficients in subband $\mathbf{w}_i$. $\mathsf{wk}_{\mathsf{DSO}}$ denotes the "wrapped key," i.e., the resolution key of the highest resolution intended for the DSO encrypted by $\mathsf{pk}_{\mathsf{DSO}}$.

transform and decryption, and will generally adhere to the protocol. However, if a recipient sees an opportunity to get data at a higher resolution than intended by the smart meter for this recipient, it will access this information.

The data concentrator needs not be a trusted entity. It only needs to be trusted to receive data and pass on this data, optionally with rate adaption. The data concentrator is not trusted with key material.

The communication links are assumed to be insecure and at least subject to eavesdropping, possibly also subject to active attacks, such as man-in-the-middle attacks.

3) Analysis of Potential Attacks: A **semi-honest recipient** (e.g., a DSO) could try to gain access to higher resolutions than the sender intended. There are different options to try: (i) derive a resolution key for a higher resolution than the wrapped key, (ii) derive higher resolutions from the lower resolutions, (iii) break the symmetric encryption of the coefficients for the higher resolutions, (iv) break the asymmetric encryption of other wrapped keys transmitted with the ciphertext (intended for other recipients). Option (i) is infeasible if a proper one-way hash function (e.g., SHA-3) is used to create the hierarchical keys. Option (ii) is infeasible given certain assumptions, which will be discussed in detail in Section *Privacy Analysis* below. Option (iii) and (iv) are infeasible if state-of-the-art cryptographic methods are used with suitable keys. In our tests we use AES and RSA. It can be concluded that even for a semi-honest recipient, the protocol achieves the intended effect of only granting access to a certain resolution.

A **collusion of multiple semi-honest recipients** will yield the plaintext data of the highest resolution granted to this group of recipients (by sharing the plaintext). However, the collusion

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

will not yield the plaintext consumption data of higher resolutions.

A **semi-honest data concentrator** trying to gain access to the data has less options than the semi-hones recipient, as it does not have any access to the keys and therefore, other than the recipient, cannot decrypt even parts of the data. Options (i) and (ii) are therefore not applicable. Options (iii) and (iv) are infeasible, if proper ciphers are used. The only information that the data concentrator has, is the number of recipients (through the number of wrapped keys), the addresses of recipients for which the data concentrator acts as a direct relay, and the number of resolutions contained in the message.

A **malicious data concentrator** can refuse to pass on the message and can thus realize a denial of service attack. It can also send bogus messages: the data concentrator can make up consumption data, acquire the public key of the DSO (or any other recipient) and then run the protocol steps for encryption (lines 1-5 in Figure 5). This can be counteracted by integrity checks. A malicious data concentrator could also perform a man-in-the-middle attack. This can be counteracted by putting authentication into place. One attack by a malicious concentrator remains that cannot be counteracted: a malicious concentrator can always decrease the resolution of the ciphertext it passes on. Recipients behind the malicious concentrator would then receive lower resolutions than intended. This circumstance would be noted by the recipient (because the available resolutions would not match the resolution implied by the included wrapped key).

Eavesdropping attacks on the communication links are not feasible as only attack options (iii) and (iv) are available, which are both infeasible. An active **malicious attack on the communication links** (such as a man in the middle attack) can realize the same attacks as a malicious data concentrator.

4) Privacy Analysis: The effectiveness of reducing personal information through the reduction in resolution is one of the central questions in evaluating the usefulness of the proposed scheme. Furthermore, as discussed in the previous section, legitimate, but semi-honest, recipients could try to obtain higher resolution information from the lower resolutions.

a) Information reduction through subsampling: For a user-centric privacy approach, it is essential to answer the question, how much privacy is introduced by repeatedly halving the sampling rate – i.e., how does a decrease in resolution actually impact the degree of personal information contained in the underlying data? And, on a related note, how much useful information is contained in a certain resolution for realizing a use case desired by the consumer (such as energy usage optimization).

It is evident that the low-passing filtering achieved by the Wavelet transform will reduce information leakage for privacy-sensitive series of load measurements. As the approach proposed here supports a high number of different resolution, it is safe to state that it will effectively increase privacy. The question remains, what target resolution the end-consumer should aim for in a given use-case.

As an example, the study of Eibl and Engel [4] explicitly aimed at tackling the relation between NILM accuracy and data resolution (Table I). Empirical material to assess the utility of a reduction of resolution for increasing privacy is provided. The four columns on the right-hand side of Table I shows the effect of decreasing resolution with the wavelet-based approach on the detection accuracy for the activities "cooking", "bathroom activities", "housework", and "presence/absence". It can be seen that decreasing the resolution is a measure to prevent detection of these activities (given the detection accuracy of current NILM methods). It should be noted that the transition from detectable to undetectable is not hard and slightly smoother than suggested by Table I.

In [4], an explicit algorithm is used for a privacy attack. Although it can be argued that typical NILM algorithms based on differences of power values should suffer from a decreased resolution this does not necessarily have to be true in general if other kinds of attacks are considered. For example the argument does not hold for approaches like the ones in [35] and [36] where the occupancy information is retrieved based on absolute values that can even be averaged over several hours.

As a candidate for a general privacy measure, Sankar, Rajagopalan *et al*. propose an information-theoretic approach that uses mutual information (MI) to evaluate privacy in [37] and [38]. From the theoretical side, there is also a relation between differential privacy and MI [39]. As a big drawback, MI has not been applied to real world data in these publications. One of the reasons for that lack may be the fact that MI is hard to estimate. There are approaches in the (Non-Intrusive Load Leveling) NILL community that use MI estimated by binning as a method to assess the similarity of the original and the changed signal [40]. The current method of choice for the estimation of MI is based on k nearest neighbors [41], [42] which has shown to be superior to the binning method which heavily depends on the bin size. In an attempt to evaluate privacy using MI, we programmed the algorithm in Matlab and successfully tested the correctness of our implementation for the correlated Gaussian example. However, the application to the real world data showed a big dependence on the number of nearest neighbors k making the results doubtful. In [41] it is mentioned that the estimation algorithm fails, if distributions are strongly peaked which is also the case here. Using ranks instead of the absolute values did not improve the estimation, so why this algorithm did not work stays an area for future research.

Instead, the regression approach of [43] and [44] is adopted. Instead of using Pearson's coefficient of correlation the more robust Spearman correlation coefficient r_{Spearman} is used here. The correlation coefficient was computed between the original sample and a wavelet approximation for each scale. In [43] and [44] this is done for finite differences of the load profile corresponding to the privacy attacks on turn-on and turn-off events. In order to account for other attacks that are based on absolute values [35], [36] this is also done for absolute values. In order to get the sign right, i.e., a measure for privacy and not for correlation $1-r_{\text{Spearman}}$ is used as a privacy measure.

In Figure 6 the result is illustrated for the mains signal of house 1 of the REDD dataset [25]. As expected the privacy

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

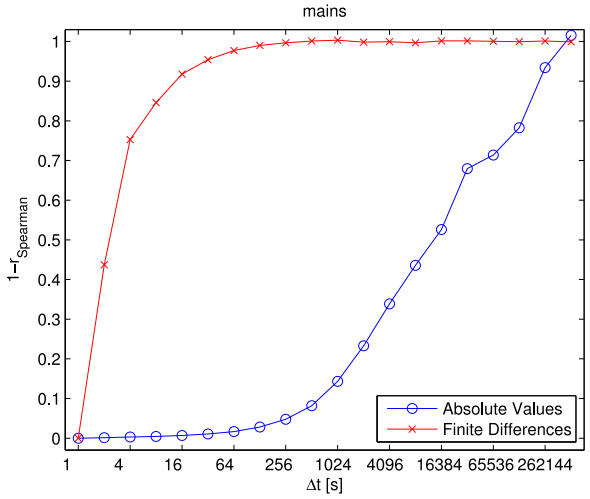


Fig. 6. Dependence of privacy, measured by 1 minus the Spearman correlation coefficient between the signal with highest resolution and lower resolutions.

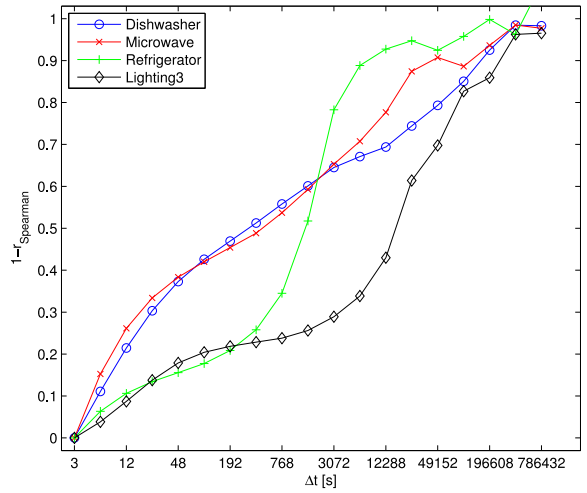


Fig. 7. Dependence of privacy, measured by 1 minus the Spearman correlation, on resolution based on absolute values for different appliances.

measure clearly increases when the resolution (here, given as a time scale) is decreased. This holds especially for the difference of subsequent values suggesting that methods based on absolute values like in [35] and [36] might be more robust with respect to a decrease of resolution.

To get further insights, the privacy measure is also computed for signals of individual appliances (Figure 7). Although absolute values are used here in contrast to [4], the results are qualitatively consistent with the results of [4]. At the same resolution, privacy of lights is estimated lower than privacy of other appliances. This is plausible due to the higher on-times of the lights [4]. Due to the more regular time-behavior of its load curve the refrigerator (typically considered as privacy-irrelevant) has a rather steep privacy increase in the middle.

b) Inferring higher resolutions from lower resolutions: Assume Eve has been granted access to resolution r of a bitstream containing a maximum resolution of d. Eve could try to use the coefficients from the lower resolutions, for which she has access, to extrapolate the higher resolution. The feasibility of this attack depends on the characteristics of the original load profile. The question is directly related to the question, how much privacy is introduced by halving the sampling rate. If the original has a high number of high frequency components (which are also crucial in NILM accuracy), significant data loss occurs going from d to r and Eve will be unable to make any assumptions on resolution d from r.

V. Combination With Other Privacy-Enhancing Technologies

The multi-resolution approach to privacy presented above is compatible with many other privacy-enhancing technologies (PETs). By combining these PETs with multi-resolution analysis, additional degrees of freedom and a broader range of choices for the end-user can be realized. In the following, we review the compatibility of the multi-resolution approach proposed here with privacy-preserving protocols found in the literature.

A. Secure Aggregation With Homomorphic Encryption

In [7], Engel and Eibl have shown that multi-resolution analysis can be used within privacy preserving protocols which directly rely on the homomorphic encryption property for secure aggregation, such as proposed by Li *et al.* [21] or Erkin and Tsudik [23]. In particular, it has been shown that when homomorphic encryption is applied to a signal represented in the wavelet domain, homomorphic additivity is not only preserved, but can be separately exploited for each resolution.

B. Additive Secret Sharing

The method proposed by Garcia and Jacobs [22] combines Paillier's homomorphic encryption with additive secret sharing. Generally, additive masking terms need no adjustment since they cancel out in the decryption step before the inverse transformation takes place. Thus, the method is compatible with the wavelet transformation. In [15], Kursawe *et al.* describe four different protocols which rely on masking. These protocols can be categorized into so-called aggregation and comparison protocols. All of the protocols are designed as simple as possible to be feasible for use in the field. All of the aggregation protocols are compatible with wavelets, and masking can be applied to each resolution separately. However, in the comparison protocols, the transformed sum of the values is in the exponent of the generating element of the Diffie-Hellman group. As the reverse transformation cannot be calculated for terms in the exponent, wavelets are not compatible with these comparison protocols.

C. Differential Privacy

In [20], Ács and Castelluccia use the modulo operation for homomorphic encryption instead of Paillier's homomorphic encryption scheme. Privacy and also confidentiality with the aggregator is achieved by masking. As stated above, the additive masking terms need no adjustment. The second main

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

feature is the addition of Laplacian noise for differential privacy. This step needs some modification to be used with the wavelet transform: Basically, the noise needs to be wavelet transformed before being added to the wavelet subbands in order to limit its impact upon reverse transformation at the recipient. The detailed mechanics of this process are out of scope of this paper and remain a subject for future work.

D. Data Integrity

The method in [45] extends [21] by preserving data integrity. The wavelet transformation is compatible with this method since it is mostly based on the ciphertext. There, it is irrelevant if the encrypted message is in its original or in a transformed form. Decryption is only done in the incremental verification process where the compatibility can be verified for each individual step.

Summarizing, the wavelet method is compatible with existing privacy preserving protocols except comparison protocols. Adaptations are needed for differential privacy.

VI. Conclusion

We have proposed a method for user-centric smart meter privacy, which uses the wavelet transform to generate a cascade of different resolutions from the load data created by a smart meter. Through the use of hierarchical keying schemes, the user can efficiently grant or deny access to external parties. Adaptation of resolution, i.e., reduction of data, can be done after encryption, also by parties lacking the keys, such as data concentrators.

The computational demands for the proposed scheme are low and make the approach feasible in an economic sense. The discussed proof of concept implementation was tested on relatively inexpensive hardware, for real-world use, significantly cheaper hardware could be used.

Wavelet-based multi-resolution privacy is compatible with many of the other PETs, which have previously been proposed in literature. We have discussed the compatibility of the proposed approach with different types of methods on a theoretical level.

In future work, the question should be addressed, how to communicate the trade-off between privacy and utility to the user. Applying the information-theoretic framework introduced by Sankar *et al.* [37] to assess privacy and utility at each resolution could be an interesting direction.

References

[1] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992.

[2] M. A. Lisovich, D. K. Mulligan, and S. B. Wicker, "Inferring personal information from demand-response systems," *IEEE Security Privacy*, vol. 8, no. 1, pp. 11–20, Jan./Feb. 2010.

[3] U. Greveler, B. Justus, and D. Löhr, "Multimedia content identification through smart meter power usage profiles," in *Proc. Int. Conf. Inf. Knowl. Eng. (IKE)*, Las Vegas, NV, USA, 2012, pp. 1–8.

[4] G. Eibl and D. Engel, "Influence of data granularity on smart meter privacy," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 930–939, Mar. 2015.

[5] D. Engel, "Conditional access smart meter privacy based on multiresolution wavelet analysis," in *Proc. 4th Int. Symp. Appl. Sci. Biomed. Commun. Technol.*, Barcelona, Spain, 2011, pp. 1–5.

[6] D. Engel, "Wavelet-based load profile representation for smart meter privacy," in *Proc. IEEE PES Innov. Smart Grid Technol. (ISGT)*, Washington, DC, USA, 2013, pp. 1–6.

[7] D. Engel and G. Eibl, "Multi-resolution load curve representation with privacy-preserving aggregation," in *Proc. IEEE Innov. Smart Grid Technol. (ISGT)*, Lyngby, Denmark, 2013, pp. 1–5.

[8] M. Jawurek, F. Kerschbaum, and G. Danezis, "Privacy technologies for smart grids—A survey of options," Microsoft Res., Cambridge, U.K., Tech. Rep. MSR-TR-2012-119, 2012.

[9] D. Engel, "Privacy-preserving smart metering: Methods and applicability (invited talk)," in *Proc. 4th Workshop Commun. Energy Syst.*, Vienna, Austria, Sep. 2013, pp. 9–16.

[10] Z. Erkin, J. R. Troncoso-Pastoriza, R. L. Lagendijk, and F. Perez-Gonzalez, "Privacy-preserving data aggregation in smart metering systems: An overview," *IEEE Signal Process. Mag.*, vol. 30, no. 2, pp. 75–86, Mar. 2013.

[11] S. Finster and I. Baumgart, "Privacy-aware smart metering: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1732–1745, Sep. 2014.

[12] M. Jawurek, M. Johns, and K. Rieck, "Smart metering depseudonymization," in *Proc. 27th Annu. Comput. Security Appl. Conf.*, Orlando, FL, USA, 2011, pp. 227–236.

[13] J.-M. Bohli, C. Sorge, and O. Ugus, "A privacy model for smart metering," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, Capetown, South Africa, 2010, pp. 1–5.

[14] C. Efthymiou and G. Kalogridis, "Smart grid privacy via anonymization of smart metering data," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Gaithersburg, MD, USA, 2010, pp. 238–243.

[15] K. Kursawe, G. Danezis, and M. Kohlweiss, "Privacy-friendly aggregation for the smart grid," in *Proc. Privacy Enhanc. Technol. Symp.*, Waterloo, ON, Canada, 2011, pp. 175–191.

[16] B. Defend and K. Kursawe, "Implementation of privacy-friendly aggregation for the smart grid," in *Proc. 1st ACM Workshop Smart Energy Grid Security (SEGS)*, Berlin, Germany, 2013, pp. 65–74.

[17] G. Danezis, C. Fournet, M. Kohlweiss, and S. Zanella-Béguelin, "Smart meter aggregation via secret-sharing," in *Proc. 1st ACM Workshop Smart Energy Grid Security (SEGS)*, Berlin, Germany, 2013, pp. 75–80.

[18] C. Dwork, "Differential privacy: A survey of results," in *Theory and Applications of Models of Computation* (LNCS 4978), M. Agrawal, D. Du, Z. Duan, and A. Li, Eds. Berlin, Germany: Springer-Verlag, 2008, pp. 1–19.

[19] E. Shi, R. Chow, T.-H. H. Chan, D. Song, and E. Rieffel, "Privacy-preserving aggregation of time-series data," in *Proc. NDSS Symp.*, 2011, pp. 1–17.

[20] G. Ács and C. Castelluccia, "I have a DREAM! (differentially private smart metering)," in *Proc. Inf. Hiding Conf.*, Prague, Czech Republic, 2011, pp. 118–132.

[21] F. Li, B. Luo, and P. Liu, "Secure information aggregation for smart grids using homomorphic encryption," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Gaithersburg, MD, USA, 2010, pp. 327–332.

[22] F. D. Garcia and B. Jacobs, "Privacy-friendly energy-metering via homomorphic encryption," in *Security and Trust Management* (LNCS 6710), J. Cuellar, J. Lopez, G. Barthe, and A. Pretschner, Eds. Berlin, Germany: Springer-Verlag, 2011, pp. 226–238.

[23] Z. Erkin and G. Tsudik, "Private computation of spatial and temporal power consumption with smart meters," in *Proc. 10th Int. Conf. Appl. Cryptography Netw. Security (ACNS)*, Singapore, 2012, pp. 561–577.

[24] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin, "Private memoirs of a smart meter," in *Proc. 2nd ACM Workshop Embedded Sensing Syst. Energy Efficiency Building (BuildSys)*, Zürich, Switzerland, 2010, pp. 61–66.

[25] J. Kolter and M. J. Johnson, "REDD: A public data set for energy disaggregation research," in *Proc. Workshop Data Mining Appl. Sustain. (SIGKDD)*, San Diego, CA, USA, Aug. 2011, pp. 1–6.

[26] H. Livani and C. Y. Evrenosoglu, "A machine learning and wavelet-based fault location method for hybrid transmission lines," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 51–59, Jan. 2014.

[27] J. Khan, S. M. A. Bhuiyan, G. Murphy, and M. Arline, "Embedded zerotree wavelet based data denoising and compression for smart grid," *IEEE Trans. Ind. Appl.*, vol. 51, no. 5, pp. 4190–4200, Sep./Oct. 2015.

[28] J. Ning, J. Wang, W. Gao, and C. Liu, "A wavelet-based data compression technique for smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 1, pp. 212–218, Mar. 2011.

[29] W. Sweldens, "The lifting scheme: A construction of second generation wavelets," *SIAM J. Math. Anal.*, vol. 29, no. 2, pp. 511–546, 1998.

[30] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Appl.*, vol. 4, no. 3, pp. 247–269, 1998.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[31] F. Knirsch, D. Engel, M. Frincu, and V. Prasanna, "Model based assessment for balancing privacy requirements and operational capabilities in the smart grid," in *Proc. 6th Conf. Innov. Smart Grid Technol. (ISGT)*, Washington, DC, USA, 2015, pp. 1–5.

[32] C. D. Peer, D. Engel, and S. B. Wicker, "Hierarchical key management for multi-resolution load data representation," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Venice, Italy, Nov. 2014, pp. 926–932.

[33] S. Imaizumi, N. Aoki, H. Kobayashi, and H. Kiya, "Hierarchical key assignment scheme for multimedia access control with modified hash chain," in *Proc. 8th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process. (IIH-MSP)*, Piraeus, Greece, 2012, pp. 293–296.

[34] L. Lamport, "Password authentication with insecure communication," *Commun. ACM*, vol. 24, no. 11, pp. 770–772, Nov. 1981.

[35] D. Chen, S. Barker, A. Subbaswamy, D. Irwin, and P. Shenoy, "Non-intrusive occupancy monitoring using smart meters," in *Proc. 5th ACM Workshop Embedded Syst. Energy Efficiency Build. (BuildSys)*, Rome, Italy, 2013, pp. 1–8.

[36] D. Chen, D. Irwin, P. Shenoy, and J. Albrecht, "Combined heat and privacy: Preventing occupancy detection from smart meters," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Budapest, Hungary, 2014, pp. 208–215.

[37] L. Sankar, S. R. Rajagopalan, S. Mohajer, and H. V. Poor, "Smart meter privacy: A theoretical framework," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 837–846, Jun. 2013.

[38] S. R. Rajagopalan, L. Sankar, S. Mohajer, and H. V. Poor, "Smart meter privacy: A utility-privacy framework," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Brussels, Belgium, 2011, pp. 190–195.

[39] W. Wang, L. Ying, and J. Zhang, "On the relation between identifiability, differential privacy and mutual-information privacy," in *Proc. 52nd Annu. Allerton Conf. Commun. Control Comput.*, Monticello, IL, USA, 2014, pp. 1086–1092.

[40] W. Yang, N. Li, and Y. Qi, "Minimizing private data disclosures in the smart grid," in *Proc. ACM Conf. Comput. Commun. Security (CCS)*, Raleigh, NC, USA, 2012, pp. 415–427.

[41] A. Kraskov, H. Stögbauer, and P. Grassberger, "Estimating mutual information," *Phys. Rev. E*, vol. 69, no. 6, 2004, Art. ID 066138.

[42] A. Papana and D. Kugiumtzis, "Evaluation of Mutual Information Estimators for Time Series", *Int. J. Bifurcation Chaos*, vol. 19, no. 12, pp. 4197–4215, 2009.

[43] G. Kalogridis, R. Cepeda, S. Z. Denic, T. Lewis, and C. Efthymiou, "ElecPrivacy: Evaluating the privacy protection of electricity management algorithms," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 750–758, Dec. 2011.

[44] G. Kalogridis and S. Z. Denic, "Data mining and privacy of personal behaviour types in smart grid," in *Proc. IEEE 11th Int. Conf. Data Min. Workshops (ICDMW)*, Vancouver, BC, Canada, 2011, pp. 636–642.

[45] F. Li and B. Luo, "Preserving data integrity for smart grid data aggregation," in *Proc. IEEE 3rd Int. Conf. Smart Grid Commun. (SmartGridComm)*, Tainan, Taiwan, 2012, pp. 366–371.

Dominik Engel (S'06–M'08) received the Ph.D. degree in computer science from the University of Salzburg, Salzburg, Austria, in 2008.

He was a Researcher with the University of Bremen, Bremen, Germany, and the University of Salzburg, and a Product Manager with Sony DADC, Anif, Austria, where he was responsible for video content security. He is a Professor with the Salzburg University of Applied Sciences, Austria, where he is the Head of the Josef Ressel Center for User-Centric Smart Grid Privacy, Security, and Control. His current research interests include smart grid security and privacy, and methods for enhancing trustworthiness of technical systems in general.

Günther Eibl (M'13) received the Ph.D. degree in mathematics and the M.Sc. degree in physics from the University of Innsbruck, Innsbruck, Austria, in 1997 and 2002, respectively.

He is a Research Associate with the Josef Ressel Center for User-Centric Smart Grid Privacy, Security, and Control, Salzburg University of Applied Sciences, Austria. He previously held research positions with the Institutes of Biostatistics, Innsbruck, and the Institute of Theoretical Physics, Innsbruck. In academic and nonacademic research, he worked in such fields as data mining and machine learning, particle and fluid simulations, computer vision, robot kinematics, control, and cryptography. His research interests include extraction of information from data with a focus on statistical modeling, data mining, and privacy preserving technologies.