JOURNAL OF OPTOELECTRONICS LASER ISSN: 1005-0086

Volume 41 Issue 8, 2022

A Lightweight Based Data Aggregation Scheme for Smart Grid Power Systems

Philani Khumalo*[1] & Bakhe Nleya[2]

*[*1&2]Department of Electronic and Computer Engineering, Durban University of Technology, Durban, South Africa*

ABSTRACT

To accomplish data aggregation securely and efficiently, it is necessary to design a scheme that is low in both computational as well as communication overheads. Thus in this paper, we propose and analyze a novel secured data aggregation scheme that ensures both privacy preservation as well as data integrity. The scheme is centered on forecasting power consumption demands for a particular neighborhood, and overall, because most attacks occur, during the transmission of data across the ICT subsystem, it thus focuses on limiting that. It does so by first forecasting its demands, and only links with the utility operator when adjustments have become necessary. The scheme utilizes a lightweight efficient noninteractive authentication mechanism in the generation and sharing of session keys. Overall, both the security analysis and performance evaluation demonstrate its efficacy in guaranteeing both privacy and security in addition to minimizing computational and communication overheads.

Keywords: lightweight encryption, smart grid, NTRU cryptosystem, computational load, communication overheads.

I. INTRODUCTION

Growing energy demands together with the urge to supply available power in a reliable, as well as efficient manner, have led to the gradual upgrading and modernizing of existing power grid systems into Smart Grids (SGs) by way of incorporating supporting information and communication technology (ICT) subsystems. The latter facilitates the two-way flow of both energy (power) and information related to the grid's performance, as well as the end user's requirements. Notably, the ICT subsystem enables key entities such as generation, distribution, transmission, and end-user subsystems to interrelate in real-time, and in the process, this achieves a well reliable, robust as well as efficiently managed SG system. The interactions of the various entities constituting the grid result in the emergence of various services and applications exchanging data throughout the interconnected systems. Whereas the SG is quite efficient in rendering its services, it, however, is exposed to various cyber security threats by adversaries. Notably, security threats vary depending on the applications. On the user end networks, the mandatory aggregation of power consumption as well as exchange of power consumption-related information n individual household area networks (HANs) or among HANs and utility's control canter (CC) can result in adversaries tempering the processes. In particular key security concerns are that during these operations, individuals' privacy, as well as aggregated data integrity, can be compromised as a result of attacks. The resource-constrained nature of associated devices, objects, and elements of the SG at the user side networks and in the SG core, in general, brings about challenges in implementing robust security measures that inevitably involve the performing of complex crypto-operations. For this reason, any measures in the form of schemes and mechanisms implemented to preserve security and privacy must be lightweight, i.e. they should minimize the generation of computational and communication overheads during operations. Nevertheless, the SG cyber-attack surface has expanded thus necessitating data security automation. In this regard, the adoption of multi-layered as well as multi-factor authentication to enhance both security and privacy is necessary. Similarly, the adoption of new cybersecurity technology stack trends means will serve as an impetus and general guidance on the architecture framework needed to secure both privacy and security. Further challenges are in that some SG elements and devices are mobile and hence this necessitates mobile software security enhancements. Periodic cybersecurity awareness training will ensure that both manufacturers and utility operators operate at the same pace and direction in combating security threats and vulnerabilities in modern SGs.

In light of what has been outlined, this work mainly addresses the security and privacy concerns within the ICT subsystem's architectures. On the customers' side networks, data security, confidentiality, privacy, and integrity must be ensured at all times. In the grid's core, measuring and monitoring units must be protected against integrity attacks, such as false data injection (FDI) attacks.

II. RELATED WORKS

Data aggregation, anonymization, and perturbation are techniques widely implemented to preserve privacy in SGs. To adhere to total privacy preservation as well as satisfying security requirements, the techniques are further augmented with multiparty computation (MPC) or homomorphic ciphering/ deciphering. Note that the ciphering part (MPC) is based on allowing individual entities (parties) to collaboratively generate using the individually owned data, but not sharing its content with the rest of the entities (parties) involved. Thus the advantage of homomorphic encryption-based techniques is that it permits cryptographic mathematical operations on ciphered text. Nevertheless, this feature does allow entities to perform the computations but without knowledge of the data contents. With anonymization, pseudonyms are instead used rather than the true entity's In that way it becomes difficult to map an individual's real

name to the energy consumption-related data. In addition To hybrid approaches, two or more primary techniques are blended thus resulting in an even more robust technique concerning privacy preservation. In this category, we have time perturbation techniques and others. Lots of research has already been done regarding privacy-preserving schemes in SG environments and we thus henceforth review them.

With regards to aggregation-based privacy preservation schemes, the authors in [1] light-weight privacy-preserving data aggregation (LPDA) that utilizes the bilinear pairing technique as well as a one-time masking method to conceal an entity's identity, whilst at the same time maintaining a lightweight aggregation. Generally, the parties are involved: several HANs within the same BAN , BAN-Gateway and the CC. Its three-phase operations are as follows:

At the starting phase, the CC computes the mandatory bilinear together with two hash functions. It retains one key designated as the master, private, and the other is made public. Both HANs and BAN–gateway register with the CC and in the process are granted private keys, which will be used for establishing static keys for communication purposes between them.

The aggregation request phase is next. The authenticated HANs receive time-stamped aggregation request messages (regarding energy consumption) from the BAN Gateway.

3During the aggregation response phase, individual HANs in the vicinity responds by collating the energy consumption message, before sending it to the BAN gateway. Note that it masks the message using its assigned static key as well as a once-off mask. Upon receiving this message, the BAN performs the necessary verifications and authentications before relaying the same message securely to the CC. The CC will also in turn verify and authenticate the received message. The scheme overall proved effective in preventing security vulnerabilities. However, its drawback is that of key management complexities as well as high computational loads. Especially now that hop-to-hop communication is involved. In [3] a homomorphic encryption-based privacy scheme is proposed. In this case, a spanning tree is formed to acquire customer consumption data. The collector, in this case, is designated as the root and all messages exchanged between nodes are encrypted. Implementation details are as follows: In the first instance, an aggregation route connecting all SMs in the targeted area is constructed. Energy consumption data is aggregated upwardly on the tree. An individual parent SM requests and acquires data messages from its child SMs and merges them with its own. This proposed approach was proven to ensure complete confidentiality since intermediary SMs cannot read the message's content. However, forgery of data is possible as there is no proper auditing of messages. Similarly in [4] privacy-preserving aggregation (EPPA) scheme is proposed. It uses primitives such as homomorphic Paillier cryptosystems, bilinear pairing, and a dynamic increasing sequence. The scheme involves three entities: SMs, residential area gateway (GW), and the utility operator. Reference [5] proposes an efficient privacy-preserving demand response (EPPDR) Scheme. It uses similar cryptographic approaches as schemes already reviewed. In particular, its initialization phase is similar to that of the EPPA scheme. Both the utility operator and GW make use of identity-based signatures in generating private and session keys, which will then be shared non-interactively between the entities involved. However, this scheme generates new session keys after each time-out period, within the same session. Based on the brief literature survey, we carried out in this and preceding sections, in this paper we propose and analyze a lightweight-based data aggregation that ensures privacy as well as confidentiality.

III. MODEL DESCRIPTION

It is assumed that both the CCs and $BANs$ cannot temper maliciously with any data received from $HANs$. Note, however, that attackers may try to intercept the data as it is exchanged between the trusted parties.

Thus to safeguard from situations in which attackers (adversaries) may extend their actions:

- *End user 's privacy*: the end customers' personal information cannot be indulged to unauthorized users.. neither can their energy consumption data or trends be divulged to unauthorized parties. To further consolidate the end user's privacy, the scheme will not divulge the actual IDs to the CC. In other words, this knowledge of these details will only be confined to $BANs$.
- *Messages Integrity and Confidentiality*: The end customer's power usage details, trends, and associated billing must be protected from any attackers. Total integrity must be ensured. In short, any attempts or acts of malicious actions must be detected in real-time.
- *Availability*: All key entities such as the BAN servers must be available and accessible throughout. This implies that the system should be completely shielded from DoS attacks.

The proposed scheme has the following desirable objectives:

- Minimization of computational loads.
- The communications overheads must be kept at a minimum or be avoided altogether
- The scheme aims to preserve the consumers' privacy.

The proposed scheme is centered on forecasting power consumption demands for a particular neighborhood, and overall, because most attacks occur, during the transmission of data across the ICT subsystem, it thus seeks to limit that.

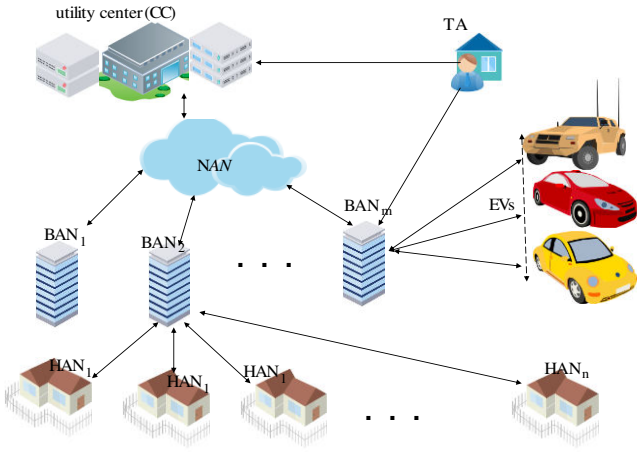


Fig. 1. : Scheme's Model Illustration

It does so by first forecasting its demands, and only links with the utility operator (CC) when adjustments have become necessary. It is a desirable goal that the scheme satisfies all desirable privacy objectives, is robust, as well as being lightweight. Furthermore, we also aim at the objective of minimizing both communication and computational overheads. As illustrated in Figure 1, $BANs$ connect to the CC via the available NAN In other words, $BANs$ do not connect directly to the CC, but via an available NAN network coverage (which is facilitated by the ICT subsystem of the SG. Note that a single SG typically comprises several $BANs = \{BAN_1, BAN_2, ...BAN_m\}$.

An individual BAN typically has limited computing resources constrained. It also interconnects several $HANs$, i.e;
$HANs = \{HAN_1, HAN_2, ...HAN_n\}$. (.1)

Note that HAN can be viewed as typically representing a standalone household and hence will comprise several household electrical appliances.
Only a single designated trusted authority (TA) assigns IDs to each SM in the SG.

NTRU Cryptosystem Review

We will rely on the NTRU, which is an open-source public-key ciphering and deciphering system that utilizes lattice-based cryptography to encrypt and decrypt session data [5].

The NTRU cryptosystem can be summarized as follows:

If n is a power of 2; and Φ has n liner factors, then $\Phi = rn+1$, $R = Z[x]/\Phi$, $q(q = 1 \mod 2n)$:

$$\Phi = \prod_{i<n} \Phi_i = \prod_{i<n} (x - \Phi_i) \mod q \quad (.2)$$

$$R_q = \frac{R}{qR} = Z[r]_q / \Phi \quad (.3)$$

where, $R_q^x \in R_q$.

In the above two equations, (2) and (3), q is a prime number.

Key generation

The key generation procedure would be as follows:

For $n, q \in Z, p \in R_q^x$, $\sigma \in R$.; if we sample the value f' from a discrete Gaussian function $D_{Z^n,\sigma}$, where $\sigma > Poly(n) * q^{0.5+\varepsilon}$, for any value of $\varepsilon > 0$, we have:

$$(sk, pk) \cup R \times R_q^x \quad (4)$$

A secret key can be generated as follows:

DOI: 10050086.2022.08.106

$f = p * f' + 1$ (5)

In the above equation $(f \mod q) \in R_q^x$, and $f = 1 \mod p$. The secret value will range from g to $D_{Z^n,\sigma}$, subject to $(g \mod q) \in R_q^x$.

We can finally recover the secret key $sk = f$ and public key $pk = h$ where;

$h = pg / f \in R_q^x$ (6)

Encryption

Given a message M, a sender S generates;

$\text{rand}.s, \varepsilon \leftarrow \overline{\text{Y}_\varepsilon}$ (7)

and ciphertext as:

$C = hs + p\varepsilon + M \in R_q$ (8)

Decryption

Upon receiving C the receiver R will decode the message using the private key f as follows:

$C' = f.C \in R_q$ (9)

$M = C' \mod p$ (10)

NTRUsign Review

The NTRUSign , also referred to as NTRU Signature Algorithm, is public-key cryptography digital signature-based and utilizes the GGH signature scheme. Its operation is mainly centered on mapping a message to a random point in a $2N$ -dimensional space, where N is defined as one of the NTRUSign parameters, and solving the closest vector problem in a lattice closely related to the NTRUEncrypt lattice.

Given N, q, d, and NB being a prime dimension, a modulus, a key size, and verification bound perimeter respectively.

Also given the existence of two polynomials f, g which are both invertible modulo q, such that their coefficients $d+1$ equal $1, d$, -1 and the remaining 0, we then have;

$h = f^{-1} * g (\text{mod } q)$ (11)

They compute polynomials (F, G) such that;

$f * G - g * F = q$ (12)

Key Generation

For an arbitrary user i we select a random polynomial $r_i \in R_q$ such that;

$f_i = f * r_i,\ g_i = g * r_i$ (13)

$F_i = F * r_i^{-1}$ (14)

$G_i = G * r_i^{-1}$ (15)

Ultimately the output is;

$Sk_i = (f_i, g_i, F_i G_i)$ (16)

Signing In Process

Upon S hashing a message M, such to create a random vector $(m_1, m_2) (\text{mod } q)$, and substituting (writing) m_1, m_2 in the following:

$G_i * m_1 - F_i \circ m_2 = A_i + q * B_i$ (17)

$-g_i \cdot m_1 + f_i * m_2 = a_i + q * b_i$ (18)

The signature on the message M is;

$s_i = f_i * B_i + F_i b_i (\text{mod } q)$ (19)

DOI: 10050086.2022.08.106

Signature Verification

The verifying entity V also hashes the received message M to create and (m_1, m_2) then calculates:

$$t_i = s_i * h(\text{mod } q) \quad (20)$$

Subject to the following:

$$\|s_i = m_1\|^2 + \|t_i - m_2\|^2 \leq NB \quad (21)$$

IV. PROPOSED SCHEME

This is a two-phase scheme, the first of which accomplishes initialization (i.e. ensuring connectivity among the different entities involved in the energy supply). The second phase addresses message exchanges within a BAN's domain.

Phase I: ***Key generation***

The key steps are summarized as follows:

The designated TA will encryption and signing in keys for both CC and BAN as follows:

For the CC's secret key f_{cc} we have;

$$f_{cc} = p * f'_{cc} + 1 \quad , f_{cc} \bmod q \in R_q^x \quad (22)$$

$$f_{cc} = 1 \bmod p \quad (23)$$

The TA also samples g_{cc} from the function $D_{Z^n,\sigma}$ such to satisfy;

$$g_{cc} \bmod q \in R_q^x \quad (24)$$

After which it calculates:

$$h_{cc} p g_{cc} / f_{cc} \in R_q^x \quad (25)$$

Thus h_{cc} is the CC á public key whereas f_{cc} is the private key.

Similarly, for the BAN gateway its keys are computed as follows:

$$f_{bam} = p * f'_{ban} + 1 \quad (26)$$

Once again $f_{cc} \bmod q \in R_q^x$ and $f_{ban} = 1 \bmod p$

The TA also samples g_{ban} from the function $D_{Z^n,\sigma}$ such to satisfy;

$$g_{ban} \bmod q \in R_q^x \quad (27)$$

After which it calculates:

$$h_{ban} p g_{ban} / f_{ban} \in R_q^x \quad (28)$$

Thus , h_{ban} is the CC is á public key whereas $f_{ban}$is the private key.

Signing keys

Once again the TA a pair of polynomials f, g invertible module q. Both satisfy $d+1$ of their roots equal 1, d roots equal -1 and the remainder equal 0. The public key for all end users is calculated according to:

$$h = f^{-1} * g(\text{mod } q) \quad (29)$$

It then computes (F, G), in which ;

$$f * G - g * F = q \quad (30)$$

In order to generate the signing key for CC, it selects $r_{cc} \in R_q$ and setting;

$$f_{ccs} = f * r_{cc}, \quad g_{ccs} = g * r_{cc} \quad (31)$$

It then further computes;

DOI: 10050086.2022.08.106

$F_{cc} = F * r_{cc}^{-1},\ G_{cc} = G * r_{cc}^{-1}$ (32)

Thus the CC 's signing keys will be;

$Sk_{cc} = (f_{cc}, g_{ccs}, F_{cc}, G_{cc})$ (33)

Correspondingly, the signing key for the BAN gateway is computed by first selecting $r_{ban} \in R_q$:

This is followed by;

$f_{bans} = f * r_{ban},\ \ g_{bans} = g * r_{ban}$ (34)

and then,

$F_{ban} = F * r_{ban}^{-1},\ G_{ban} = G * r_{ban}^{-1}$ (35)

Thus the BAN 's signing keys will be;

$Sk_{ban} = (f_{ban}, g_{bans}, F_{ban}, G_{ban})$ (36)

Generation of IDs

Each SM is assigned an d ID, $ID_1, ID_2, ...ID_n$. At regular intervals corresponding pseudo IDs are generated according to;

$ID_{new} = h(ID_{old})$ (37)

where h is a hash function.

Electricity Demand Forecast

This is done according to a forecasting function $g(\ )$ and for each HAN, the forecasted demand is;

$x_i = g(HAN_i)$ (38)

Thus for each cluster, the BAN aggregates the forecasted demands as follows

$x = \sum(x_1, x_2, ..., x_n) + \varepsilon$ (40)

Where ε denotes a backup. Note that the backup is mainly derived from EVs ;

$C_{EV} = \sum C_i\ , 1 \le i \le N_{EV-\exp ected}$ (41)

Thus we have;

$\varepsilon = r * C_{EV}$ (42)

subject to $0 < r < 1$ a scaling factor

Note that at initialization phase, an optimal number of EVs required to work as energy buffers is determined by the BAN according to:

$\min\ \ N_{EV}(m)$ (43)

Subject to:

$\varepsilon(m) \le \sum_i C_i(m),\ \ i \in \{1,..., N_{current}(m)\}$ (44)

$N_{EV}(m) \le N_{current}(m),\ \ \ N_{current}(m) \in \{1,..., N_{\max}(m)\}$ (45)

$m \in \{1,...,100\}$

Power consumption Agreement

Note that x is considered as the aggregated demand per BAN by the CC. It is never aware of each individual HAN 's requirements in this regard.

The Agreement Request message

This is an agreement between the BAN and CC. The BAN initially sends an agreement request message m_a to CC. This typically involves sending the requested (forecasted) power demand x in encrypted form i.e. x is hashed to yield $(x_1, x_2) (\mathrm{mod}\ q)$.

$G_{ban} * S_1 - F_{ban} * S_2 = A_{ban3} + qB_{ban3}$ (46)

$-g_{bans} * S_1 + f_{bans} * S_2 = a_{ban3} + q * b_{ban3}$ (47)

Thus the signature is;

$S = s_{ban3} = f_{bans} * B_{ban3} + F_{ban} * b_{ban3} (\mathrm{mod}\ q)$ (48)

This will yield S and s_{ban3}. Consequently, the BAN computes;

$$m_5 = S\|s_{ban3}\|T_{s_5}\|k_5 \quad (49)$$

After encrypting m_5 the BAN sets $s_5, \varsigma \leftarrow \overline{\gamma_\alpha}$ and also uses h_{cc} to generate:

$$m_b = h_{cc}s_5 + p\varsigma_5 + m_5 \in R_q \quad (50)$$

At the CC, f_{cc} is used to decrypt m_b. $_{ban3}$ s is also verified, so is the validity of the time stamp.
If the need arises, the BAN can adjust the requested power according to the algorithm in Figure 2.

1. BAN Electricity Share Adjustment Procedure
2. x : The xed demand for BAN
3. y : The current actual demand for BAN
4. z : The EV remaining capacity
5. $\beta : \beta = \|x - y\|$: The dierence between x and y
6. **If** $(x > y \,\&\, \beta < z)$ then;
7. $\beta \leftarrow EV_{battery}$
8. else if $(x < y\beta\hat{} > z)$ then
9. $B - z \leftarrow CC$
10. else if $(x < y \,\&\, \beta > z)$ then
11. $\phi - z \rightarrow CC$
12. end if

Fig. 2. Power renegotiation algorithm1

V. RESULTS AND DISCUSSION

The analysis of this scheme is in two parts, security analysis, and evaluation. The security analysis is done using a set of Java-based scripts of the NTRU public-key cryptosys-tem, comprising the signature NTRUSign and encryption NTRUEncrypt schemes. [6].

Security Analysis

As emphasized in the introductory stages of the subsection, the goal of the proposed scheme is to ensure privacy for the end-users, associated data (mostly billing) as well as entities. In addition, it should be able to provide sufficient confidentiality, integrity, availability, authenticity, and accountability guarantees. We thus analyze all these aspects in relation to the scheme's capabilities.

Individual End-User's Privacy Preservation. With respect to an individual's private information, the scheme thrives by all means to maintain privacy. This includes the concealment of the end-users ID, location, and power-consuming patterns. Whereas it is possible that the CC can be a point of launching attacks by adversaries, however, note that the scheme, (CC) does not have details of the end user's details. Only the BAN has the information. Neither is it able to extract finer details of individual power consumption bills, since the BAN delvers such information in aggregated form (i.e for all the customers connected to it). In any case, all these messages are in encrypted form e.g messages from BAN to CC are delivered in encrypted form. The use of both private and public keys means that only the CC can decrypt the messages, as it has a corresponding deciphering key. Messages relayed from $HANs$ to the BAN are also in encrypted form.

Messages' confidentiality is guaranteed. The agreement between parties such as the CCs and $BANs$ are concluded and exchanged using public keys. As such confidentiality is maintained between these entities. The same applies to messages between the $HANs$ and BAN gateways. As was earlier cited, the TA assigns identities to SMs and for this reason, MiTM attacks will not succeed. In short, the SM 's ID s real identity is concealed. The use of signing signatures makes it further difficult to intercept and decrypt any messages, as an only authority with the correct signatures can do so.

Integrity of Messages Exchanged. Concerning the integrity of exchanged messages between CC and BAN, note that they must be hashed and signed (using a private key). The SMs also combine energy consumption-related messages together with their IDs, prior hashing them and relaying to $BANs$. The SM details can only be validated by the BAN since it has this information stored in its database Likewise, the database information is secured since stored data is encrypted by a private key disclosed only to the TA.

DOI: 10050086.2022.08.106

Authenticity guarantees: Public keys are used to authenticate entities such as CC and BAN. In that way, their messages are also authenticated as it is them who encrypt, formulate and encrypt the messages.

Resources Availability: BAN gateways are secured from DoS attacks. The number of $HANs$ to a given BAN is strictly limited and any extras (i.e attackers) will immediately be detected

Accountability: Individuals are at liberty to validate as well as verify their bills at the local BAN. This is because the latter has information related to price changes. Moreover, there is a relatively low expected volume of message exchanges between the customer end and $BANs$. In that way adversaries may find it hard to intercept messages. The NTRU cryptosystem also prevents adversaries from obtaining any knowledge from any intercepted data. In summary, the scheme preserves customers' privacy.

Performance Evaluation

We will look at the efficacies of the scheme in the following subsections. In this regard, we only carry out an analytical evaluation.

Communication overheads

In designing protocols and schemes some of whose operational aspects involve communication, it is important to evaluate the levels of communication overheads. Communication overhead refers to the extra data bits in the headers and message trailer flags. Such extra data does assist in proper addressing, flow, and error control as well as delineation at the receiver end.

Relatively, low volumes of messages are exchanged between the various parties in the proposed protocol scheme. This is partly because most of them are first aggregated, then dispatched as a multiparty message.

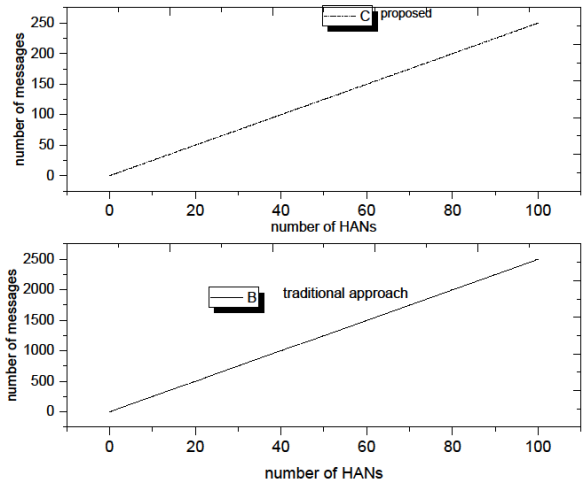


Fig. 3: Comparisons of communication overheads (a) proposed (b) traditional

E.g $BANs$ do not participate in the initialization phase. However, only two messages each are exchanged by both CC and BAN when negotiating the power-share agreement, In the second phase $HANs$ restrict themselves to sending demand messages provided this has been necessitated by a sudden change in demand or power tariffs. Figure 3 plots the communication overheads. The same Figure plots overhead levels for traditional approaches for comparison purposes. In this case, the number of $HANs$ connections per BAN is varied gradually. It is clearly shown that the proposed scheme does lower the number of messages exchanged (and consequently communication overhead) in comparison with traditional approaches.

We subsequently explore the communication overhead loads for varying numbers of demand messages., i.e the following five cases are considered:

- Case I : a fraction of the total number of $HANs$ send a single power demand message over a 24 hour period (day), whilst the rest do not send any requests at all.
- Case II: Each HAN send a single power demand message per 24 hour period (day).
- Case III: Some $HANs$ send two power demand messages each, others send a single message, and the rest do not send at all over a 24-hour period (day).

- Case IV: Each HAN send a couple of power demand messages per 24 hour period (day).

Case V: Each HAN dispatches three power demand messages every day.
Figure 4 plots the variation in communication overhead for the proposed taking into regard the five scenarios explained earlier. As can be observed from the graph, the number of messages to about 350 domain clusters has 120 $HANs$.

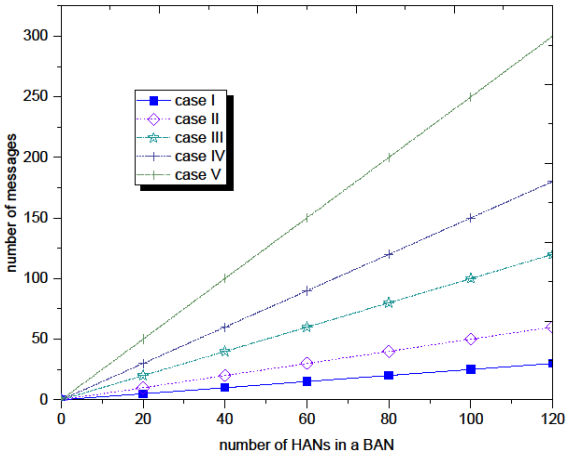


Fig. 4: Communication overhead considering various scenario cases

Computation complexity
In the context of this work, computational overhead (loading) would be a combination of excess or indirect computation time, memory, bandwidth, or related resources that are required for the proposed scheme to perform the desired tasks (focused design objectives).

In this case, the computational times for signing (T_S), verification operations (T_V), encryption (T_E), and decryption (T_D) are taken into consideration.

Note that during the initialization phase, CC and BAN each performs a single cyphering operation, single decryption, as well as a once-off signing/verification. This equates to a computational time of :

$$C_I = 2\times[T_E + T_D + T_S + T_V] \quad (51)$$

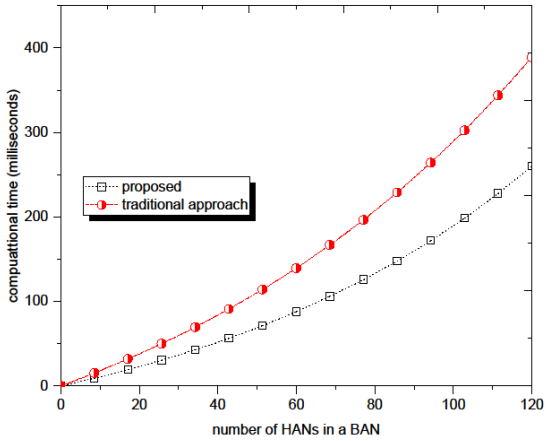
For the next phase (phase II) a HAN is likely to be involved in message exchanges in the form of demanding extra power allocations. In this case, it performs a single encryption operation which will be decrypted by the associated BAN. This equates to $T_E + T_D$ per data message. The likelihood that tariffs might change and hence necessitates communication between CC and BAN, thus the computation time is $2\times[T_S]+T_V$. If the number of $HANs$ is m, the total computational time becomes;

$$m(T_E + T_D) + (2T_S + (m+1)T_V) \quad (52)$$

In the next phase, i.e. billing, the message is sent to CC from BAN this necessitating one encryption, one decryption, one sign and one verification process. Following the method used in [150], to approximate, the aggregate computational times, we have;

$$C_{proposed} = 90\times[T_E + T_D + T_S + T_V] \quad (53)$$

$$C_{traditioal} = 810\times[T_E + T_D + T_S + T_V] \quad (54)$$



Fig, 5:Computation overhead Traditional .vs. Proposed scheme.

Figure 5 plots, the aggregate computational times of our proposed scheme. From the plotted graph, it is observed that an increase in the number of $HANs$ results in computational time increases. However, by comparison, the increases are much lower for the proposed scheme. Thus overall, our scheme manages to execute fast, despite the limited computational resources. The computational time fit within the expected time frame scales of a fully fletched SG network.

We now evaluate a worst-case scenario, in which all $HANs$ in a cluster domain send the maximum possible dumber of power demand messages. Figure 6 provides a plot of the computational times in this worst-case scenario.

Once again, the aggregated computational times for the proposed versus traditional approach schemes are plotted and compared. The proposed scheme has a much reduced computational overhead hence the computational times are comparably less.

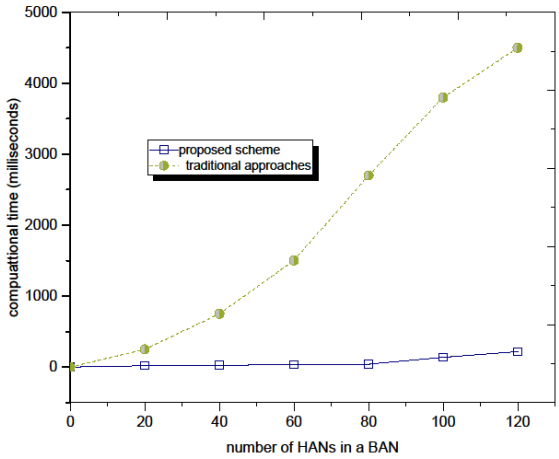


Fig. 6: Computation overhead Traditional .vs. Proposed scheme

DOI: 10050086.2022.08.106

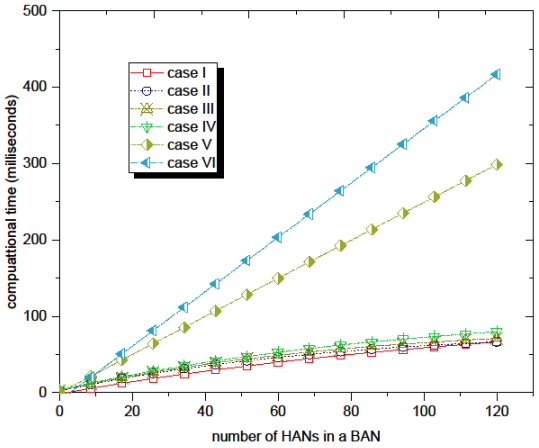


Fig. 7: Computation overheads comparisons

The five case scenarios (repeated) plus an additional sixth case, are once again considered as we explore aggregated computational times.

- Case I : a fraction of the total number of $HANs$ send a single power demand message over a 24 hour period (day), whilst the rest do not send any requests at all.
- Case II: Each HAN send a single power demand message per 24 hour period (day).
- Case III: Some $HANs$ send two power demand messages each, others send a single message, and the rest do not send at all over a 24 hour period (day).
- Case IV: Each HAN send a couple of power demand messages per 24 hour period (day).
- Case VI: Each HAN sends a maximum possible number of power demand messages per 24 hour period (day).

As can be observed from Figure 7, the computational complexity relatively rises with increases in power demand messages. We thus can conclude that the proposed scheme guarantees privacy and at the same time it minimizes computational and communication overhead levels.

VI. CONCLUSION

In this paper, a lightweight based data aggregation scheme that ensures privacy, as well as confidentiality, is proposed. It is centered on forecasting power consumption demands for a particular neighborhood, and overall, because most attacks occur, during the transmission of data across the ICT subsystem, it thus focuses on limiting that. It does so by first forecasting its demands, and only links with the utility operator (CC) when adjustments have become necessary. It is a desirable goal that the scheme satisfies all desirable privacy objectives, is robust, as well as is lightweight. Furthermore, the scheme has the objective of minimizing both communication and computational overheads.

Conflicts of Interest

None of the authors has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

Funding Statement

The work was partly funded by the Faculty of Engineering jointly with the Research Office as part of support to my PhD studies.

Acknowledgments

We acknowledge the support we got from colleagues in the Department.. The work was equally contributed by the two authors.

REFERENCES

[1] R. Lu, X. Liang, X. Li, X. Lin and X. Shen, EPPA: An Efficient and Privacy-Preserving Aggregation Scheme for Secure Smart Grid Communications,: IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 9, pp. 1621-1631, Sept. 2012, doi: 10.1109/TPDS.2012.86.

[2] H. Li, X. Lin, H. Yang, X. Liang, R. Lu and X. Shen, EPPDR: An Efficient Privacy-Preserving Demand Response Scheme with Adaptive Key Evolution in Smart Grid: IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 8, pp. 2053-2064, Aug. 2014, doi: 10.1109/TPDS.2013.124.

[3] A. Abdallah and X. Shen, Lightweight Security and Privacy Preserving Scheme for Smart Grid Customer-Side Networks: IEEE Transactions on Smart Grid, vol. 8, no. 3, pp. 1064-1074, May 2017, doi: 10.1109/TSG.2015.2463742.

[4] D. He, N. Kumar, S. Zeadally, A. Vinel and L. T. Yang, Efficient and Privacy-Preserving Data Aggregation Scheme for Smart Grid Against Internal Adversaries: IEEE Transactions on Smart Grid, vol. 8, no. 5, pp. 2411-2419, Sept. 2017, doi: 10.1109/TSG.2017.2720159.

[5] M. B. Line, I. A. Tøndel and M. G. Jaatun, Cyber security challenges in Smart Grids: 2011 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies, 2011, pp. 1-8, doi: 10.1109/ISGTEurope.2011.6162695.

[1] GitHub - tbuktu/ntru: Java implementation of NTRUEncrypt and NTRUSign