On the Digital Certificate Management in Advanced Metering Infrastructure Networks

Vassilios G. Vassilakis*, Ioannis D. Moscholios†, John S. Vardakas‡, Michael D. Logothetis§
* Dept. of Computer Science, University of York, York, United Kingdom
† Dept. of Informatics & Telecommunications, University of Peloponnese, Tripolis, Greece
‡ Iquadrat Informatica, Barcelona, Spain
§ Dept. of Electrical & Computer Engineering, University of Patras, Patras, Greece

***Abstract*—The smart grid (SG), generally referred to as the next-generation power system, is considered as a revolutionary and evolutionary regime of existing power grids. Among the emerging SG applications, the advanced metering infrastructure (AMI) enables automated, two-way communication between a smart meter (SM) and a public utility company. To authenticate a message, the sender (e.g., a SM) signs the message with its private key using a pre-defined digital signature algorithm. To verify the message, the recipient verifies the sender's certificate and then the sender's signature using the sender's public key. In some cases, however, a previously issued certificate for a network node needs to be revoked. In this paper we investigate two possible approaches for the certificate management of SMs in AMI networks. These are based on the traditional certificate revocation lists (CRLs) and on the Bloom filters. We compare the two approaches in terms of the required packet size for the distribution of the revoked certificate serial numbers. We also discuss the advantages and limitations of each approach.**

Keywords*—Certificate revocation list; advanced metering infrastructure; smart grid; Bloom filter.

I. Introduction

The smart grid (SG) [1], generally referred to as the next-generation power system, is considered as a revolutionary and evolutionary regime of existing power grids. More importantly, with the integration of advanced computing and communication technologies, the SG is expected to greatly enhance the efficiency and reliability of future power systems with renewable energy resources, as well as distributed intelligence and demand response [2].

The SG requires efficient and reliable communication networks for management and coordination [3]. Among the emerging SG applications, the advanced metering infrastructure (AMI) enables automated, two-way communication between a smart (utility) meter (SM) and a public utility company. For example, a SM may collect power consumption information from various smart home appliances, such as washing machines and refrigerators, and send the aggregated measurement data to the electric utility company. A SM typically has an IP address and may report the aggregated measurement data to the utility company. A SM may also send control commands to smart appliances to turn them on and off based on various cost optimization objectives [4]. In general, the goal of an AMI is to provide the utility companies with real-time data about power consumption and allow the customers to make informed choices about energy usage based on the price at the time of use. A typical AMI network consists of a number of SMs that are connected to a gateway, which forwards the data and control messages to/from the utility company. Our considered AMI network architecture is described in more detail in Section III.

Along with the salient features of the SG and the AMI, cyber security emerges to be a critical issue because large numbers of heterogeneous electronic devices are interconnected via communication networks throughout critical power facilities. This has an immediate impact on reliability of such a widespread infrastructure [5]–[8]. Some of the basic SG operations include establishing remote connections with sensor nodes and performing updates and configurations. These operations require the development of appropriate methods for enabling secure connection, secure boot, and secure update of SG sensor nodes. Due to the resource-constrained nature of typical SG nodes, the suitability of traditional Internet protocols, such as IPsec and TLS, needs to be re-examined.

Focusing on the AMI network, the main security requirements include identity and message authentication, message integrity, non-repudiation, accountability, and access control [11]. These requirements are typically achieved by using the public-key cryptosystems in which each SM has one private and one public key. The private keys are kept secret, whereas the public keys are announced in the network. The binding of a public key with a particular SM identity is usually performed via a public key certificate that is issued by a (trusted) certificate authority (CA). To authenticate a message, the sender SM signs the message with its private key using a pre-defined digital signature algorithm. To verify the message, the recipient (SM, gateway, or utility company) verifies the sender's certificate and then the sender's signature using the sender's public key. In some cases, however, a previously issued certificate for a network node needs to be revoked.

In this work, we investigate two possible approaches for digital certificate revocation in AMI networks, namely the certificate revocation lists (CRLs) and Bloom filters (BFs). In particular, by considering realistic scenarios we compare the two approaches in terms of the required packet size and discuss their limitations.

The rest of the paper is structured as follows. In Section II we discuss different approaches for digital certificate management. In Section III we describe our considered AMI network architecture. In Section IV we describe and analyze a BF-based approach for the management of digital certificates in AMI networks. In Section V we discuss the main limitation of the

BF-based approach, namely the false positives. In Section VI we compare the traditional CRL-based approach with the BF-based approach in terms of the required packet size. Finally, in Section VII we conclude and discuss possible future directions. Also, in Table I we present the list of abbreviations used in this paper.

II. Digital Certificate Management Approaches

A. Certificate Revocation Lists (CRLs)

As discussed in Section I, in some cases a previously issued certificate for a network node needs to be revoked. This may happen if, for example, the particular node has been captured by an adversary or the node's private key has been compromised. Hence, if some nodes show malicious behavior, their certificates must be quickly revoked in order to protect the rest of the network. Certificate revocation is typically done by the CA who regularly announces the CRLs. Hence, each network node before accepting a signed message must check with the CRL to ensure that the sender's certificate has not been revoked. In the case of the AMI network, the traditional approach that requires the CRLs to be stored in each SM and be frequently updated by the CA, introduces significant storage and communication overhead [9]

B. The Online Certificate Status Protocol (OCSP)

An alternative approach to CRLs is to use the online certificate status protocol (OCSP) [10]. Similar to the CRLs, the OCSP enables a requesting party (e.g., a SM) to determine the revocation state of a certificate. However, this information is stored in a remote server rather than in the SM itself. When a CA signs a certificate, it will typically include an OCSP server address in the certificate. Hence, when a node (e.g., a SM) is presented with another node's certificate, it will send a query to the OCSP server. The latter listens to queries and responds with the revocation status of the certificate. The main advantage of the OCSP over the CRL approach is that since an OCSP response contains less information than a typical CRL, it puts less burden on network and client resources. On the other hand, a big disadvantage of the OCSP approach is that the certificate verification cannot be done locally and there is a need to deploy new devices to act as OCSP servers. Furthermore, it is insecure to install the OCSP servers on gateways that are typically deployed in streets and lack physical security. Finally, it is not efficient and scalable to use the CA as the OCSP server due to significant communication delays and overhead for performing the queries from remote SMs [11].

C. Bloom filter (BF)-Based Certificate Management

Due to aforementioned limitations of the traditional CRLs and the OCSP, attempts have been made to mitigate their performance and scalability limitations using BFs [9], [11], [12]. A BF is a space-efficient probabilistic data structure that is used to concisely represent a set and allows highly efficient set membership queries [13]. For this reason, the BF-based approach appears as an attractive solution for enabling an efficient management of digital certificates in AMI networks. In particular, the set of revoked certificates is encoded by the CA into a BF and is sent to all SMs. Each SM stores the BF locally. When the certificate verification is required, the SM performs a membership test to determine whether a given certificate is in the BF or not. Compared to the traditional CRL-based approach, the BF-based approach requires less storage space. Compared to the OCSP-based approach, the BF-based approach does not require sending queries to a remote server and, hence, is faster and more bandwidth-efficient. However, the BF-based approach has some probability of giving false positives in a membership test. That is, an element (i.e., a certificate in our case) may appear to belong to the set (i.e., the set of revoked certificates in our case) when in fact it does not. This is a direct consequence of the compressed nature of BFs.

TABLE I. List of Abbreviations

AMI	Advanced Metering Infrastructure
BF	Bloom Filter
CA	Certificate Authority
CRL	Certificate Revocation List
CSN	Certificate Serial Number
GW	Gateway
HAN	Home Area Network
OCSP	Online Certificate Status Protocol
NAN	Neighborhood Area Network
SHA	Smart Home Appliance
SG	Smart Grid
SM	Smart Meter
WAN	Wide Area Network

False positives during the certificate verification could be very problematic, because they may cause a node to discard a legitimate message by wrongly assuming that the sender's certificate has been revoked. Fortunately, as it is shown in Section VI, the false positives probability can be kept low by increasing the size of the BF. Furthermore, various BF optimizations can be preformed to further decrease the false positives rate [14]. On the other hand, a big advantage is that the BF-based approach has no false negatives. That is, if a certificate is not found in the BF, this means that it has not been revoked and the receiving node can trust the message signature.

III. Advanced Metering Infrastructure (AMI) Network Architecture

In this Section we present our considered AMI network architecture. The architecture is shown in Fig. 1 and comprises three tiers:

- Home area network (HAN): It consists of one SM and several smart home appliances (SHAs), such as TV, refrigerator, security cameras, etc. The network topology is typically a star with the SM as the central node. The communication of the SM and SHAs is done using some appropriate two-way short-range communication technology, such as ZigBee/IEEE 802.15.4.
- Neighborhood area network (NAN): It consists of one gateway (GW) and several SMs. These typically form a wireless mesh network and use some appropriate medium-range communication technology, such as WiFi/IEEE 802.11.
- Wide area network (WAN): Is used to connect the GW to the electric utility and the CA. This typically requires a long-range wired or wireless communications technology, such as cellular LTE-based network.

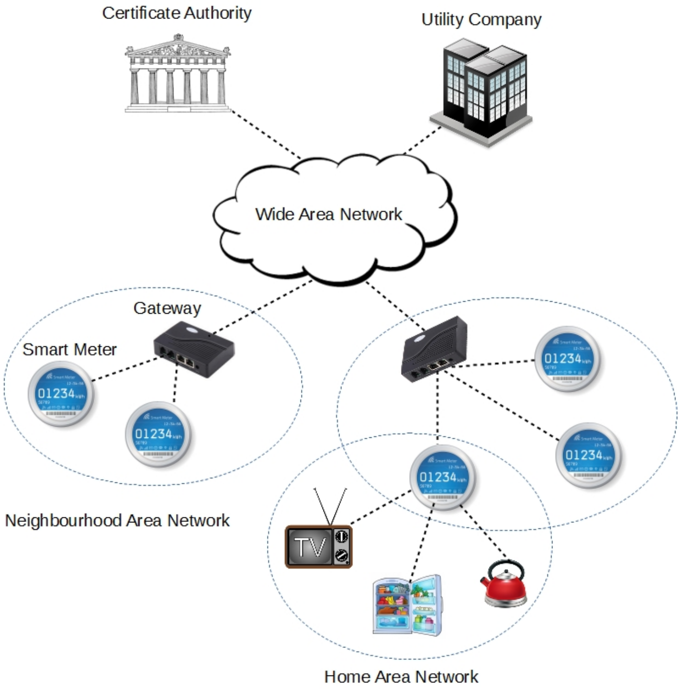

Fig. 1. Advanced metering infrastructure (AMI) network architecture.

IV. Bloom Filter Based Digital Certificate Management

Consider a CA that needs to disseminate the CRL to the SMs in a NAN. The number of certificates in the CRL is denoted by N. A certificate is included in the CRL by adding the certificate's serial number (CSN) which is unique for a given CA. The CRL also includes the CA's signature. Hence, the traditional CRL has a length of:

$$L_{CRL} = L_{CSN} N + L_{CAS} \quad (1)$$

where L_{CSN} is the length of a CSN and L_{CAS} is the length of CA's signature. A typical CSN has a length of 16-20 bytes, depending on the CA, and the CA's signature is about 700 bytes [12].

An alternative approach is to encode the revoked CSNs into a BF. The basic concept of the BF-based approach is illustrated in Fig. 2. A BF is a bit-vector that is used to store elements of a set (e.g., the revoked CSNs in our case). The length of the BF is denoted by m. It is possible to add elements into the BF, but it is not possible to remove elements.

A. Adding Elements into a BF

Assume that initially the BF is empty. This is denoted by setting all bits to 0. To add an element into the BF, the element is hashed by K hash functions, $h_1(), h_2(), \ldots, h_K()$. Each hash function returns an integer in the range $[1, \ldots, m]$. This integer represents the position of the m-bit vector that will be set to 1. The result of these K hashing operations is that some bits in the BF are set to 1. In Fig. 2 we show an example of adding two CSNs into an m-bit BF using $K = 3$ hash functions. For instance, we have $h_1(CSN_1) = 2$. This means that after applying the 1st hash function to the 1st element, the position 2 in the BF will be set to 1. Similarly, we have $h_3(CSN_2) = 6$. This means that after applying the 3rd hash function to the 2nd element, the position 6 in the BF will be set to 1.

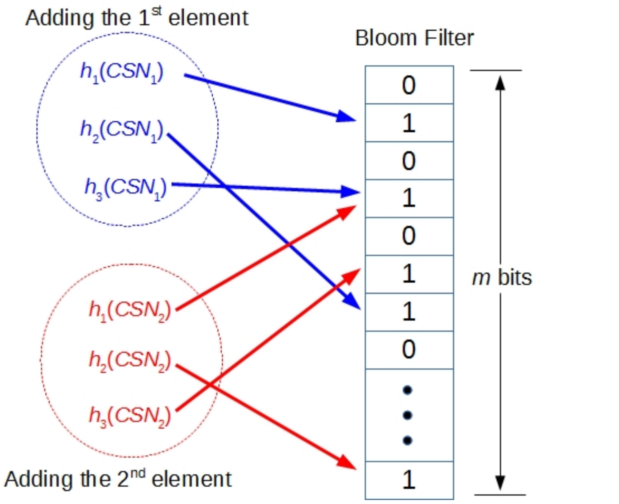

Fig. 2. The Bloom filter concept: adding two elements. $K = 3$ hash functions.

B. Testing Element Membership in a BF

Consider now a non-empty BF that contains (all) the elements of a set. To test whether an element is a member of the set or not, the element is hashed using the K aforementioned hash functions. Then the resultant K bit positions are checked against the corresponding bit positions in the BF. If all these bit positions in the BF are 1, then the test result is positive. That is, the element is considered to be a member of the set (although false positives may occur, as will be discussed in Section V, below). If at least one of the K positions in the BF is not 1, then the test result is negative. That is, the element is not a member of the set (recall that there are no false negatives in BF membership tests). As an illustration of the membership test, consider the BF of Fig. 2. To perform the test on CSN_1, we determine the $K = 3$ hash values, which give us: $h_1(CSN_1) = 2$, $h_2(CSN_1) = 4$, and $h_3(CSN_1) = 7$. After that, we verify that all three given positions, 2, 4, and 7, in the BF are set to 1. Hence, the set membership of the element CSN_1 has been verified.

V. False Positives in Bloom Filters

A. An Example

As mentioned in Subsection IV-A an element is added into the BF by setting appropriate bit positions to 1. These bit positions are returned by K hash functions. However, it may happen that by adding some elements to the BF, all K bit positions that correspond to another element (that was not intended to be added) are set to 1. One such case is shown in Fig. 3. In this example, the BF that was created by the adding elements CSN_1 and CSN_2 of the example of Section IV. Assume that we would like to perform the membership test for another element, CSN_3. As illustrated in Subsection IV-B, the test is performed by hashing the element with K hash functions. If the resultant hash values are $h_1(CSN_3) = 6$,

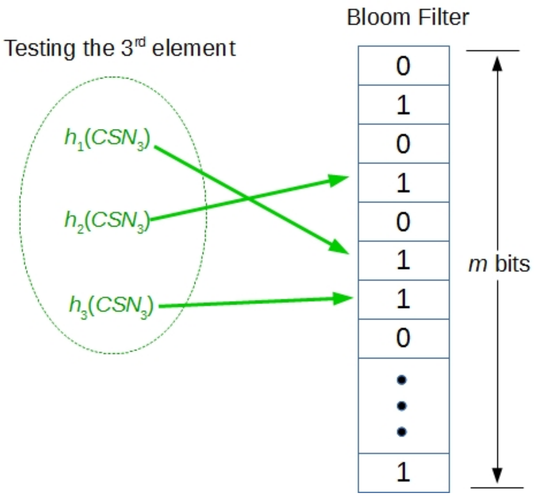

Fig. 3. Illustrating the false positives during a membership test.

$h_2(CSN_3) = 4$, and $h_3(CSN_3) = 7$, then the test will give a positive result since all three resultant bits positions (6,4, and 7) are set to 1 in the BF. However, since CSN_3 is actually not a member of the set, we refer to this test result as a false positive.

B. Calculating the False Positive Probability

The false positive probability of a BF can be determined as follows. Consider an m-bit vector $\boldsymbol{b} = (b_1, b_2, \ldots, b_m)$ representing a BF. The BF is initially empty, that is $b_i = 0, \forall i \in [1, \ldots, m]$. Assume that we add N CSNs, denoted as $CSN_j, j \in [1, \ldots, N]$, into the BF using K independent hash functions, denoted as $h_k(), k \in [1, \ldots, K]$. Let us denote by $n_{k,j}$ the hash value that results after hashing CSN_j with the k-th hash function, $h_k()$. For instance, in Fig. 2 we have $n_{1,1} = h_1(CSN_1) = 2$ and $n_{2,3} = h_2(CSN_3) = 6$.

When applying the 1st hash function to the 1st CSN, the hash value $n_{1,1}$ denotes the bit position in the vector $\boldsymbol{b}$ that will be set to 1. Since $\boldsymbol{b}$ has m bits, the probability that any particular bit b_i is set to 1 is $1/m$. Consequently, the probability that any particular bit b_i remains 0 is $1 - 1/m$. When applying K hash functions to the 1st CSN, due to the independence of hash functions, the probability that a bit b_i remains 0 is $(1-1/m)^K$. Similarly, when N CSNs are hashed with K hash functions, the probability that a bit b_i is 0 is $\Pr(b_i = 0) = (1 - 1/m)^{KN}$. Consequently, the probability that a bit b_i is set to 1 is:

$$\Pr(b_i = 1) = 1 - (1 - \frac{1}{m})^{KN} \tag{2}$$

When performing a membership test for a particular CSN CSN_j, order to have a false positive, all hash values $n_{1,j}, n_{2,j}, \ldots, n_{K,j}$ must point to the bit positions that are set to 1 in the vector $\boldsymbol{b}$. This happens with probability:

$$P_{FP}(CSN_j) = \prod_{k=1}^{K} \Pr(b_{n_{k,j}} = 1) = (1-(1-\frac{1}{m})^{KN})^K \tag{3}$$

By introducing the approximation $(1 - 1/m)^{KN} \approx \mathrm{e}^{-KN/m}$, which holds as $m \to \infty$, (3) can be re-written as:

$$P_{FP} = (1 - (1 - \frac{1}{m})^{KN})^K = (1 - \mathrm{e}^{-KN/m})^K \tag{4}$$

It can be shown that P_{FP} is minimized when the (optimal) number of hash functions is [14]:

$$K_{opt} = \frac{m}{N} \ln 2 \tag{5}$$

Hence, substituting (5) into (3) we get:

$$P_{FP} = (1 - \mathrm{e}^{-\ln 2})^{(\ln 2)m/N} = 0.5^{(\ln 2)m/N} \tag{6}$$

The above can be rewritten as:

$$\ln P_{FP} = -\frac{m}{N}(\ln 2)^2 \tag{7}$$

Finally, the above can be rewritten as follows to express the required BF size, m, in terms of the false positive probability, P_{FP}, and the number of elements, N, in the set:

$$m = -\frac{N \ln P_{FP}}{(\ln 2)^2} \tag{8}$$

VI. Evaluation

In this Section we compare the BF-based certificate management with the traditional CRL-based approach. The comparison is made in terms of the required packet size. We assume that both the BF and the CRL are created by the CA and are sent to the SMs at regular intervals (e.g., daily).

In Fig. 4 we present the required packet size versus different numbers of revoked certificates, N, for both approaches. In the case of the CRL-based approach, we consider two different sizes for the CSNs: $L_{CSN} = 15$ bytes and $L_{CSN} = 20$ bytes. The packet size corresponds to the CRL length of (1) and the CA's signature length is chosen to be $L_{CAS} = 700$ bytes.

In the case of the BF-based approach, the CSN size does not affect the packet size. This is because each CSN is hashed into K bit positions (as mentioned in Section IV). Furthermore, since the BF-based approach introduces false positives during the membership test, we distinguish three different cases for the false positive probability: 1%, 0.5%, and 0.1%. The packet size corresponds to the BF size of (8) but is converted from bits to bytes in order to enable the comparison with the CRL-based approach.

The results of Fig. 4 show that in all cases the BF-based approach requires a significantly smaller packet size compared to the CRL-based approach. This results in the advantages in reducing both the storage overhead at the SMs as well the communication overhead between the CA and the SMs. Although the BF-based approach suffers from the false positives, we observe that the packet size reduction is significant even in the case that we require a very low false positive rate of 0.1%.

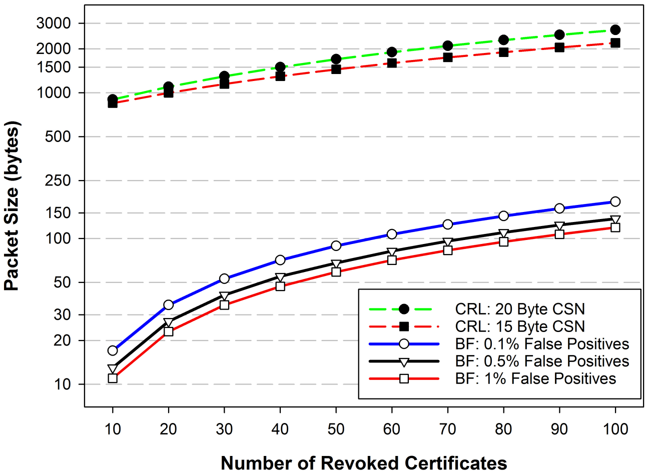

Fig. 4. Packet size versus the number of revoked certificates for CRL-based and BF-based approaches.

VII. Conclusions and Future Work

In this paper we consider an AMI network architecture and investigate possible approaches for the digital certificate management. In particular, we explore two solutions for communicating the revoked certificate serial numbers to the smart meters, namely the traditional CRL-based approach and the BF-based approach. Analytical calculations based on realistic parameters show that the BF-based approach requires much lower packet size compared to the CRL-based approach for the same number of revoked certificates. On the other hand, the BF-based approach suffers from the false positives. That is, there is some probability that a genuine certificate may appear as revoked during the certificate verification process. However, as shown, the false positive rate can be reduced to some acceptable level by increasing the packet size. In our considered scenario, even when the upper limit for false positive is as low as 0.1%, the BF-based approach still results in much smaller packet size compared to the CRL-based approach. In our future work we plan to further investigate the applicability of BFs for certificate management in AMI networks and to perform comparisons with the online protocols, such as the OCSP.

References

[1] X. Fang, S. Misra, G. Xue, and D. Yang, “Smart grid - The new and improved power grid: A survey,” IEEE Communications Surveys & Tutorials, vol. 14, no. 4, 2012, pp. 944-980.

[2] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, “A survey on demand response programs in smart grids: Pricing methods and optimization algorithms,” IEEE Communications Surveys & Tutorials, vol. 17, no. 1, 2015, pp. 152-178.

[3] N. Saputro, K. Akkaya, and S. Uludag, “A survey of routing protocols for smart grid communications,” Computer Networks, vol. 56, no. 11, 2012, pp. 2742-2771.

[4] S. Tennina, D. Xenakis, M. Boschi, M. Di Renzo, F. Graziosi, and C. Verikoukis, “A Modular and flexible network architecture for smart grids,” Proc. International Conference on Ad-Hoc Networks and Wireless, 2015, pp. 273-287.

[5] P. Yi, T. Zhu, Q. Zhang, Y. Wu, and L. Pan, “Puppet attack: A denial of service attack in advanced metering infrastructure network,” Journal of Network and Computer Applications, vol. 59, 2016, pp. 325-332.

[6] M. Nabeel, X. Ding, S.-H. Seo, and E. Bertino, “Scalable end-to-end security for advanced metering infrastructures,” Information Systems, vol. 53, 2015, pp. 213-223.

[7] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, “A Survey on cyber security for smart grid communications,” IEEE Communications Surveys & Tutorials, vol. 14, no. 4, pp. 2012, pp. 998-1010.

[8] W. Wang and Z. Lu, “Cyber security in the smart grid: Survey and challenges,” Computer Networks, vol. 57, no. 5, 2013, pp. 1344-1371.

[9] M. Mahmoud, K. Akkaya, K. Rabieh, and S. Tonyali, “An efficient certificate revocation scheme for large-scale AMI networks,” Proc. IEEE International Performance Computing and Communications Conference (IPCCC), 2014 , pp. 1-8.

[10] J. J. Haas, Y.-C. Hu, and K. P. Laberteaux, “Efficient certificate revocation list organization and distribution,” IEEE Journal on Selected Areas in Communications, vol. 29, no. 3, 2011, pp. 595-604.

[11] K. Rabieh, M. Mahmoud, and S. Tonyali, “Scalable certificate revocation schemes for smart grid ami networks using bloom filters,” IEEE Transactions on Dependable and Secure Computing, 2015.

[12] A. Alrawais, A. Alhothaily, C. Hu, and X. Cheng, “Fog computing for the internet of things: Security and privacy issues,” IEEE Internet Computing, vol. 21, no. 2, 2017, pp. 34-42.

[13] B. H. Bloom, “Space/time trade-offs in hash coding with allowable errors,” Communications of the ACM, vol. 13, no. 7, 1970, pp. 422-426.

[14] L. Carrea, A. Vernitski, and M. Reed, “Optimized hash for network path encoding with minimized false positives,” Computer Networks, vol. 58, 2014, pp. 180-191.