Private inter-network routing for Wireless Sensor Networks and the Internet of Things

Paolo Palmieri  
Cranfield University  
Centre for Electronic Warfare  
Information and Cyber  
Shrivenham, Swindon SN6 8LA, UK  
paolo.palmieri@cranfield.ac.uk

Luca Calderoni  
University of Bologna  
Dept. of Computer Science and Engineering  
Cesena 47521, Italy  
luca.calderoni@unibo.it

Dario Maio  
University of Bologna  
Dept. of Computer Science and Engineering  
Cesena 47521, Italy  
dario.maio@unibo.it

ABSTRACT
As computing becomes increasingly pervasive, different heterogeneous networks are connected and integrated. This is especially true in the Internet of Things (IoT) and Wireless Sensor Networks (WSN) settings. However, as different networks managed by different parties and with different security requirements are integrated, security becomes a primary concern. WSN nodes, in particular, are often deployed “in the open”, where a potential attacker can gain physical access to the device. As nodes can be deployed in hostile or difficult scenarios, such as military battlefields or disaster recovery settings, it is crucial to avoid escalation from successful attacks on a single node to the whole network, and from there to other connected networks. It is therefore crucial to secure the communication within the WSN, and in particular, maintain context information, such as the network topology and the location and identity of base stations (which collect data gathered by the sensors) private.

In this paper, we propose a protocol achieving anonymous routing between different interconnected IoT or WSN networks, based on the Spatial Bloom Filter (SBF) data structure. The protocol enables communications between the nodes through the use of anonymous identifiers, thus hiding the location and identity of the nodes within the network. The proposed routing strategy preserves context privacy, and prevents adversaries from learning the network structure and topology, as routing information is encrypted using a homomorphic encryption scheme, and computed only in the encrypted domain. Preserving context privacy is crucial in preventing adversaries from gaining valuable network information from a successful attack on a single node of the network, and reduces the potential for attack escalation.

CCS CONCEPTS
•Security and privacy → Network security; Embedded systems security;

KEYWORDS

ACM Reference format:

1 INTRODUCTION
Attacks against smart devices, sensor networks and the Internet of Things (IoT) are increasing in both frequency and magnitude. In particular, malware, intended as malicious software or hardware, poses a significant security threat. As recently discovered by Kryptowire, a company operating in the cybersecurity field, more than 700 million phones, cars and other smart devices running the Android operating system were compromised and equipped with malware capable of stealing every kind of information the device was able to deal with (including files stored on the device or the messages sent from the device to one another) [10]. The malware, allegedly developed by Shanghai Adups Technology Company, sent massive amounts of sensitive data about the devices and their users’ activities back to servers in China. This latest example proves that commonly adopted business models in smart devices industry are inadequate and potentially dangerous, as the devices are usually not engineered following an effective security strategy. As such, they are a preferred target for cybercrime groups, which can exploit their ubiquity to build botnets or, in case they are deployed as part of critical or sensitive infrastructure, to compromise the regular workflow of the network and infrastructure itself [17].

The emerging security threat, however, is not slowing the growing diffusion of systems and services based on IoT and heterogeneous sensor networks, propelled by the relentless advances in the production of low-cost embedded devices and sensors. As these technologies are usually deployed in wireless environments, Wireless Sensor Networks (WSN) have become a suitable solution for an increasing number of applications, including health monitoring, smart agriculture, weather sensing, intrusion detection applications and industrial control [7, 12]. In urban and suburban contexts, these networks are often connected one to each other, enabling management control over complex scenarios. However, in spite of the extensive research in the area, the Internet of Things and in particular the interconnection of WSNs still face many security and privacy challenges [11].
The wireless nature of the communication link makes the network inherently vulnerable to eavesdropping. Moreover, IoT and WSN nodes are often deployed in unsecured areas or outdoor, where they can be subject to tampering, and a potential attacker might be able to gain control of one or more nodes. For this reason, the security of the network should be preserved even in the presence of internal adversaries. In particular, both the communication between the nodes and the context information (including the location of the nodes and the network organization) should be protected [11].

In order to preserve the privacy of the communication, nodes can employ encrypted communication protocols when feasible, due to the low-power nature and limited computational capabilities of most devices. With regard to context privacy, instead, the primary aim is to hide the location and identity of the nodes in the network, as well as the overall network structure and topology [3]. This is crucial especially in WSNs, which are in general highly vulnerable to attacks targeted at base stations (the nodes collecting the data gathered by the sensors). In fact, failure of a base station can disrupt network operation, making it an ideal target for an attacker. In order to prevent adversaries from launching both remote, software-based attacks and physical attacks the location of base stations and the network topology should be therefore concealed [4]. A basic strategy to achieve this is flooding and transmissions of fake or dummy packets, which make network traffic observation more difficult [19]. More complex strategies include the use of random walks to route packets anonymously [9]. Random walks have been adopted in a number of designs: Zhang proposed self-adjusting directed random walks in [20], while GROW (Greedy Random Walk) [18] introduced a two-way random walk, from both source and destination, that can reduce the chance of an eavesdropper being able to collect location information. Finally, layers of encryption can be used to protect the information at each hop in the walk [5]. More recently, advanced anonymity techniques have been applied to IoT and WSN, and in particular onion routing protocols derived from Tor [14].

1.1 Contribution

In this paper we introduce a novel anonymous routing mechanism, based on the Spatial Bloom Filter data structure and homomorphic encryption. The proposed construction is targeted at preserving context privacy within a network composed of a number of interconnected subnetworks. In particular, our construction can find direct application in all the settings where different networks, such as wireless sensor networks or networks of smart or embedded devices, are connected to form a larger network. The anonymous routing mechanism achieves the following goals: encrypt communication between nodes; hide the identity and location of the sending and receiving nodes in a communication between two different subnetworks; hide the network structure and topology to all the nodes; and hide the origin and destination of any communication between subnetworks to the routing layer (that is, the network infrastructure that connects the different subnetworks and is responsible for the routing of packets between them). These properties enable context privacy and security against adversaries who control one or more nodes within the network, and prevent attacks aimed at taking over control of the network.

2 PRELIMINARIES

In the following we present the main building blocks of the proposed routing mechanism: first, the Spatial Bloom Filter (SBF) [2, 15]. Second, the homomorphic encryption operations that make it possible to compute the SBF in the encrypted domain. For the latter, we base our construction on the Paillier cryptosystem [13], although any equivalent alternative cipher may be used.

2.1 Spatial Bloom Filters

A Bloom Filter (BF) is a data structure that represents a set of elements in a space-efficient manner [1]. Bloom filters are widely used in networking protocols, and have a variety of network security applications [8]. Recently, Calderoni, et al. proposed a compact data structure based on Bloom filters, designed to store location information [2, 15]. The structure, called Spatial Bloom Filter (SBF), was originally designed for location privacy applications. Two private positioning protocols were proposed with the SBF, both aimed at

![Figure 1: Sets $\Lambda_1$, $\Lambda_2$ and $\Lambda_3$ (representing three subnetworks) are used to construct a SBF. Three hash functions are used to map each element into the filter. In the first step of this example, the identifiers of two nodes belonging to $\Lambda_1$ are processed by the hash functions, resulting in the value ‘1’ being written six times into the SBF. The construction proceeds likewise for elements of $\Lambda_2$ and $\Lambda_3$. Two kinds of collisions are possible, as highlighted: the first is intra set; the second takes place when elements of sets marked with a greater label overwrite those with a lower value. The probability of both events can be controlled to prevent false positives.](image-url)
keeping both the user’s exact position and the provider’s monitored areas private. The SBF was recently evaluated in a comparative assessment with other similar privacy-preserving techniques, showing promising properties in several domains [16]. In particular, the SBF is suitable for application beyond the location privacy field. In this paper, we use the SBF to build a novel private routing protocol for interconnected networks, a typical scenario in the IoT and distributed sensor networks domain. In the following, we briefly review the data structure and its properties relevant to the proposed construction; a full definition and discussion can be found in [2, 15].

A Spatial Bloom Filter extends the original Bloom Filter idea in order to support several sets composed of elements belonging to a specific domain \( \mathcal{E} \). A SBF can be used to perform membership queries on the originating set of elements without knowledge of the set itself but, contrary to the BF, a SBF can be constructed over multiple sets. Querying a Spatial Bloom Filter for an element returns the identifier of the specific set among all the originating sets in which the element is contained, minus a false positive probability. The SBFs hold several probabilistic properties useful to control the false positive probability throughout the originating sets.

An SBF can be defined as follows: let \( \mathcal{E} \) be a domain specific set of elements (in this paper elements represent the IDs of network nodes) and let \( S = \{ \Delta_1, \Delta_2, \ldots, \Delta_z \} \) be a set of subsets such that \( \Delta_1 \subseteq \mathcal{E} \) and \( S \) is a partition of the union set \( \Delta = \bigcup_{\Delta_i \in S} \Delta_i \). Let \( O \) be the strict total order over \( S \) for which \( \Delta_i < \Delta_j \) for \( i < j \). We define the Spatial Bloom Filter over \( (S, O) \) as the set of pairs

\[
B^S(S, O) = \bigcup_{h \in H} \{(h(\delta), i) \mid \delta \in S, h \in H, s.t. \delta \in \Delta_i \land h(\delta) < h(\delta') \land h(\delta), i < j \},
\]

where \( H = \{ h_1, \ldots, h_k \} \) is a set of \( k \) hash functions such that each \( h_i \in H : \{0,1\}^* \to \{1, \ldots, m\} \), that is, the hash functions take binary strings as input and output a number uniformly chosen in \( \{1, \ldots, m\} \).

A spatial Bloom filter \( B^S(S, O) \) can be represented as a vector \( b^S \) composed of \( m \) values, where the \( i \)-th value

\[
b^S[i] = \begin{cases} 1 & \text{if } (i, i) \in B^S(S, O) \\ 0 & \text{if } (i, i) \notin B^S(S, O) \end{cases}.
\]

In the following, when referring to a SBF, we refer to its vector representation \( b^S \).

The construction of an SBF starts by setting all values in \( b^S \) to 0. Then, starting from the first set \( \Delta_1 \), each element belonging to the set is processed by each function \( h \in H \). Let us suppose \( h(\delta) = i \); in that case, the \( i \)-th value of \( b^S \) will be set to 1 (as 1 is the label associated to \( \Delta_1 \)). Elements belonging to subsequent sets \( \Delta_2, \ldots, \Delta_z \) are processed likewise. It is important to note that collisions between two distinct values are subject to the SBF collision rule: labels with higher value overwrite those with lesser value. This procedure is exemplified in Figure 1.

In order to check whether or not an element \( \delta_u \) is member of the set \( \Delta_1 \in S \), two conditions need to be met:

\[
\exists h \in H : b^S[h(\delta_u)] = i \quad \text{and} \quad \forall h \in H, b^S[h(\delta_u)] \leq i.
\]

Substantially, one single \( b^S[h(\delta_u)] = 0 \) is sufficient to state that \( \delta_u \) is not a member of \( \Delta_1 \). On the contrary, if \( b^S[h(\delta_u)] \neq 0 \) for each hash function, then \( \delta_u \) is a member of the set \( \Delta_1 \) minus a false positive probability; \( i \) is the lesser value among those returned by the set of hash functions.

### 2.2 Homomorphic encryption

The Paillier cryptosystem [13] is an asymmetric encryption scheme featuring notable homomorphic properties. Specifically, in this paper we rely on the additive homomorphism of the Paillier encryption function over \( \mathbb{Z}_n \), which leads to several identities, among which we recall:

\[
\forall m \in \mathbb{Z}_n, k \in \mathbb{N} : \ D(E(m)^k \mod n^2) = km \mod n \quad (4)
\]

This multiplicative property ensures that an encrypted plaintext raised to the power of a constant \( k \) will decrypt to the product of the plaintext and \( k \).

In the proposed protocol, we apply this multiplicative scheme on a vector basis, achieving a secure entrywise product (also known as Hadamard product). We refer to this operation as to Private Hadamard Product, and we represent it with \( \odot \).

We note here that the Paillier cryptosystem may not be suitable for some heavily computationally constrained devices: however, the proposed protocol can be achieved over any additively homomorphic cipher.

### 3 A SECURE ROUTING STRATEGY

We study a setting where different, heterogeneous subnetworks are interconnected, creating a larger network. The subnetworks are connected to each other by the routing layer, that is, the part of the overarching network infrastructure that manages and routes inter-network communication. Each subnetwork is composed of multiple nodes, and can be connected to the routing layer either directly, or through one or more gateways. In the case of Wireless Sensor Networks, these gateways could also represent the base

---

**Figure 2:** A sample sensor network composed of three subnetworks \( \Delta_1, \Delta_2 \) and \( \Delta_3 \). Each subnetwork, composed by a set of nodes, represents an Area of Interest (AOI) as described in [2, 15], and is marked with a label. Anonymous routing of packets between the subnetworks (done by the routing layer) is achieved using an SBF representing the network.
stations (where information from the sensor node is collected). The aim of our construction is to enable private routing between the subnetworks. In particular, we want to prevent an attacker that controls one or more nodes of the network from being able to learn the topology and structure of the network. Specifically, he should not be able to: determine the number of subnetworks, other than those where he controls a node; the location of any node in the network, that is, to which subnetwork a node belongs. We define the security of our construction as follows:

**Security Definition.** Private routing between different subnetworks in a wider network is achieved when: any node in the network only needs the ID’s of other nodes in order to communicate with them, and learns nothing about their position within the network; any subnetwork gateway only routes packets transparently between the subnetwork and the routing layer, and, similarly to other nodes in the subnetwork, learns nothing about the positions of nodes outside its subnetwork.

The security of the construction is analysed in Section 4.

### 3.1 Routing strategy

Each node of the network is identified by a unique, random ID. Contrary to the IP address, the ID does not contain or imply any information regarding the network structure. Within the network, nodes communicate using their respective IDs, following a tunneling and encapsulation strategy for lower level protocols (such as TCP/IP) similar to the one used in other private-preserving protocols, including onion routing [6]. In practice, communication between nodes of the network is first tunneled to the local gateway, then from the gateway to the routing layer, from then to the destination gateway and finally to the destination node. Gateways do not have an active role, and they only relay communication between the nodes in their subnetwork and the routing layer transparently. In general, each party in the communication will not reveal unnecessary information to the following one. The gateway of the sending node, in particular, will not communicate the ID of the node to the routing layer. As the receiving gateway does not know to which node in its subnetwork the communication is destined to, it broadcasts the packets to all nodes in the receiving subnetwork. Since communication is encrypted (as explained in the following), only the intended receiver will be able to decrypt the information. An example of network structure is presented in Figure 2.

### 3.2 Packets and routing information

Messages transmitted through the network using the anonymous routing protocol are composed of two parts: a header, which contains routing information; and a payload, which is encrypted and encapsulates the communication being anonymously routed (in practice, the payload contain packets of lower layer protocols such as UDP or TCP).

In order to encrypt the payload, we assume that each node in the network has a public/private key pair, and a key distribution mechanism exists between the nodes, so that each node knowing another node’s ID either knows or can retrieve the node’s public key as well. Encryption of the payload is performed by the sending node $s$ using the public key $P_k_r$ of the receiving node $r$, which can then decrypt the transmission using its secret key $S_k_r$. As communication is routed anonymously, the ID of the sender is included in the encrypted payload as well, in order for the receiving node to be able to respond.

The use of random IDs to identify the nodes removes the need to know the destination IP address in order to initiate communication, and hides the originating IP. It also means that no communication is possible without knowledge of the ID of the destination node. However, in order for the routing to be anonymous, the header does not include the ID of the sending and receiving nodes, but only routing information in the form of an homomorphically encrypted SBF. In particular, the network maintainer builds an SBF representing all the nodes in the network and their respective subnetwork. As shown in Figure 1, the elements of the set over which the SBF is built are the IDs of the nodes, while the sets are the subnetworks, each represented by a label. The SBF built this way, $b^s$, is encrypted using a homomorphic encryption scheme, as explained in Section 2.1. In this construction we use the Paillier cryptosystem [13], but any cipher with equivalent homomorphic properties can be used. In particular, other more lightweight cryptosystems could be more suitable for resource-constrained devices. The secret key $S_k$ of the homomorphic key pair is known by the routing layer, while the public key $P_k$ and the encrypted filter $Enc_{P_k_r}(b^s)$ are distributed to all the nodes. The nodes also know the set of hash functions used in constructing the filter.

Table 1 summarizes the information that each party in the protocol needs in order to communicate. The information is divided in two sets: information related to the encryption mechanism (such as public keys), in the upper row; and information related to network communication (including IDs and IP addresses), in the lower row. In this paper, we assume that knowledge of the ID of a node equates to knowledge of its public key: any suitable key distribution scheme can be applied to achieve this.

### 3.3 Routing protocol

In the following, we describe the communication between a sender node $s$ and a receiver node $r$ in two different subnetworks ($A_s$ and $A_r$ respectively) over the anonymous routing protocol. This process is visible in Figure 3.

Communication happens as follows:

<table>
<thead>
<tr>
<th>Node $j$</th>
<th>Routing Layer</th>
<th>Network Maintainer</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ID_j$</td>
<td>$P_k_j$</td>
<td>$PK, S_k$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$ID_i$</td>
<td>$P_k_i$</td>
<td>$PK, S_k$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subnetwork IP</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>122.200.64/24</td>
<td>1</td>
</tr>
<tr>
<td>122.200.43/24</td>
<td>$k$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Node IP</th>
<th>Node ID</th>
<th>Area</th>
<th>Key pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>$IP_j$</td>
<td>$ID_j$</td>
<td>1</td>
<td>$PK, S_k$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$IP_i$</td>
<td>$ID_i$</td>
<td>$k$</td>
<td>$PK, S_k$</td>
</tr>
</tbody>
</table>

Table 1: Information available to each stakeholder. The first row identifies cryptographic keys owned by the stakeholder and information related to the filter; the second row routing information and IDs of the nodes in the network.
Figure 3: Operation of the private routing protocol. Node \( s \) wants to securely transmit message \( msg \) to node \( r \). Node \( r \) belongs to subnetwork \( i \), but \( s \) only knows \( r \)'s ID \( (ID_r) \). Communication proceeds as follows: \( s \) generates the SBF related to \( ID_r \) and counts the number \( z \) of non-zero values in it; the filter is then multiplied (through an homomorphic encryption operation) by the shared encrypted filter \( b^r \). The resulting filter is then sent to the routing layer, together with \( z \). The routing layer decrypts it, and computes the destination subnetwork \( i \). The payload (that is, the encrypted message) is then routed to the subnetwork \( i \) and node \( r \), either through a gateway or by broadcast. \( r \) receives the message and decrypts it.

1. The sender node \( s \) identifies the anonymous identifier \( ID_r \) of the receiving node \( r \). \( s \) then builds an SBF with \( ID_r \) as only element, using the set of hash functions and counting the number \( z \) of non-zero values in the resulting filter. Then, the node multiplies the filter it just built by the encrypted filter \( Enc_{pk^e}(b^r) \), using the multiplicative properties of the cryptosystem. We call the resulting combined encrypted filter \( e^z \). The sender shuffles \( e^z \), and sends it to the gateway, with \( z \) and the encrypted payload \( Enc_{pk_r}(msg) \).

2. The sending gateway relays transparently the information received by \( s \) to the routing layer.

3. The routing layer decrypts \( e^z \): the decrypted filter is composed of zeros, and a number of non-zero values \( i \). If the number of \( i \)'s is equal to \( z \), then the receiving node \( r \) exists. The value \( i \) identifies the correct subnetwork to which the communication will be routed. In case of differing values, the smallest is used (see Section 2 for an explanation). Finally, the routing layer transmits the encrypted payload \( Enc_{pk_r}(msg) \) to the correct subnetwork \( \Lambda_i \).

4. The gateway of \( \Lambda_i \) receives the encrypted payload and broadcasts it to all the nodes in the subnetwork.

5. The intended receiver \( r \) receives \( Enc_{pk_r}(msg) \) and decrypts it using its secret key \( Sk_r \).

The properties of the Spatial Bloom Filters introduce the possibility of false positives, with two possible scenarios: first, an element outside the sets over which the filter has been built could be recognized as member of a set; second, an element that is a member of a set \( X \) could be recognized as member of set \( Y \). The former case has no real implications for the proposed protocol: it would only apply to the case of a node in the network using non-existing or unknown IDs. But as no public key is associated to these IDs, communication is impossible. The latter case could result in the wrong routing being applied to the communication: however, we note that the probability of this event can be calibrated through the use of appropriate parameters (such as the length of the filter and the number of hash functions) during the filter construction, and a filter can be tested after it has been built to verify that no false positive (in the sense of wrong recognition) is possible.

4 SECURITY ANALYSIS

In order to analyse the security of our construction, we discuss three separate scenarios: in the first, an attacker gains control over a node in the network; in the second, the attacker controls a subnetwork gateway, and in the third, the attacker controls the routing layer (or part of it). In all three cases, we assume the attacker will not actively disrupt network traffic, but will limit himself to observing traffic visible to him in order to learn information on the network structure and topology (context information). This is called a semi-honest behaviour. In the following, we show how in each of the three cases the attacker is unable to learn any meaningful information on the network structure, and therefore the security definition is satisfied. Security cannot be guaranteed in case the attacker controls simultaneously 1) the routing layer and 2) either one or more nodes, or one or more gateways, or a combination of the two. The extent to which security is compromised in this case depends on the number of nodes and gateways controlled, and is limited to the parts of the network the attacker has visibility of.

**Attacker controlling a node.** In this case, the attacker can read all information sent and received by the node, and learns the IDs of all the other nodes with which the controlled node can communicate. The attacker also learns the encrypted filter, but has no information to decrypt it. The attacker cannot learn the IP addresses corresponding to the nodes, as they are unknown to the controlled node and cannot be derived from the respective IDs. Similarly, the attacker cannot learn the network structure (the position of the nodes within the subnetworks and the number of subnetworks), as the routing of sent and received packets is achieved anonymously.

**Attacker controlling a subnetwork gateway.** An attacker controlling a gateway will learn all the identity of all the nodes in the respective subnetwork. However, he will not be able to read any information sent and received by the nodes, as the payloads are
encrypted. Similarly, he will not learn the destination of sent packets or the origin of received ones, as the routing information $e_r$ is encrypted. Finally, the attacker cannot learn the network structure as per the case above.

**Attacker controlling the routing layer.** In this case, the attacker will be able to watch the flow of information between the different subnetworks. However, due to the properties of the SBF, even being able to decrypt the encrypted routing information $e_r$ will not enable him to learn the identity of the receiving node $r$. Similarly, he cannot learn the identity of the sending node $s$, as this is encrypted within the payload, and the sending gateway will not communicate it to him.

5 CONCLUSIONS

In this paper, we present a private routing protocol that can be used to communicate anonymously between different networks. Our protocol can be applied in a variety of Internet of Things scenarios: from Wireless Sensor Networks, to interconnected IoT systems composed by different devices or infrastructures.

Our protocols achieve context privacy by using homomorphic encryption, tunneling and the Spatial Bloom Filters. In particular, we achieve the following properties: communication between nodes can only be read by the intended receiver; the network structure and topology (context information) is kept private to all nodes; the identity and location of the sending and receiving nodes in two different subnetworks is kept private to the routing layer; and the routing layer is oblivious to the origin and destination of any communication between subnetworks. These properties enable context privacy and security against adversaries who control one or more nodes within the network, or even the routing layer. Therefore, the proposed anonymous routing protocol can prevent attacks aimed at taking over control of the network.

REFERENCES