

Requirements and Challenges in the Design of Privacy-aware Sensor Networks

Efthimia Aivaloglou Stefanos Gritzalis Charalabos Skianis
Information and Communication Systems Security Laboratory
Department of Information and Communication Systems Engineering
University of the Aegean, Samos, Greece
{eaiv, sgritz, cskianis}@aegean.gr

Abstract—Sensor networks are set to become a truly ubiquitous technology that will affect the lives of the people in their application environment. While providing the opportunity for sophisticated, context-aware services, at the same time sensor networks impose great privacy risks. This paper discusses privacy issues in sensor networks, by identifying the requirements for privacy preserving deployments, analysing the challenges faced when designing them, and discussing the main solutions that have been proposed.

I. INTRODUCTION

In the forthcoming era of ubiquitous computing, mobile ad hoc networking technologies and sensor networks as a specialisation of these, are expected to play an essential role. Large scale sensor networks, promising to offer increased data collection capabilities through the use of small sensing devices, are not only envisioned to be globally deployed for commercial, scientific or military purposes, but also to be integrated into our daily lives, providing context rich information for highly sophisticated services.

The use of sensor network applications, utilising numerous almost invisible sensors that constantly monitor their environment, collect, process and communicate a variety of information, inevitably causes concerns related to their potential of abuse and the risks they impose to the privacy of individuals. The potential risks are aggravated by the fact that different types of sensor networks may be deployed for different purposes, others being trusted, for example subscription based sensor networks offering health services, others being partially trusted, like those deployed for customer assistance in shopping malls, and others untrusted, like surveillance networks the users might be completely unaware of. Even sensor networks initially deployed for legitimate purposes may be abused or violated. Historically, it is believed that as surveillance technology has become cheaper and more effective, it has been increasingly used for privacy abuses [1].

The application domain of sensor networks affects both the sensitivity of the information collected and the possibility of attackers to induce information other than the data monitored. In order to preserve privacy, it primarily has to be ensured that sensed information is confined to the sensor network and is accessible only to authorized parties. However, the infrastructureless nature of sensor networks, the computational capability limitations of sensor nodes, and the fact that in-

network data aggregation operations are performed, are only some of the challenges faced. Moreover, an issue that must be taken into consideration during the design of realistic security mechanisms is that security services do not provide core network functionality, and, being supportive services, should impose reasonable computational overhead.

This paper is outlined as follows: In Section II the privacy requirements in sensor networks are identified. Section III outlines the main challenges confronted and the constraints applying on the design of security solutions. The solutions that have been proposed and correspond to each of the requirements set are then discussed in Section IV. The paper concludes with some remarks in Section V.

II. PRIVACY REQUIREMENTS

Preserving the privacy of the individuals that act within an information system generally entails two aspects: Keeping their personal information confidential and accessible only to authorised parties, and protecting their private space from undesirable interruptions provoked intentionally by system abuse. Within the scope of sensor networks, unauthorised location tracking of wearable sensor nodes would consist a possible passive privacy breach. An active violation of privacy in a health care or home security sensor network scenario could consist of masquerading and inserting false data into the network to raise false alarms.

Although specific privacy requirements depend both on the application space of a sensor network, which determines the sensitivity of the information collected and communicated, and the users' preferences, roles and level of trust to the service provider, strong privacy will be preserved by designing the sensor network so that the following can be achieved.

A. Confidentiality of the sensed data

As for all privacy preserving systems, data confidentiality must be ensured through message encryption. For sensor networks, this requirement becomes even more crucial, because an outsider can induce information by correlating the results reported from multiple sensors surrounding an individual. Moreover, because of the in-network data aggregation operations, data of different granularity -and sensitivity with respect to the user's privacy- is being communicated and needs to be protected.

B. Protection of the communications' context

Ensuring the confidentiality of the messages' content does not always suffice, since an adversary might induce sensitive information by observing the communications' contextual data, especially since they can be correlated with prior information about the people and the physical locations that are being monitored by a set of sensors. For example, the availability of both spatial and temporal data may allow tracking the relative or actual -by correlation with prior knowledge- location of the mobile sensor nodes that might be carried by users, which would constitute a serious privacy breach. The information that can thus be considered sensitive is the network identity of the communicating parties, the frequency of the communications, the traffic patterns, the size of the messages and the location and time at which the sensor's measurements are being sent.

C. Indistinguishability

The identity of the individuals acting within a sensor network must be protected from illegitimate users. Anonymity can be defined [2] as the state of not being identifiable within a set of subjects, referred to as the anonymity set, which itself may be a subset of a worldwide set of subjects who send messages. What anonymity mechanisms can ensure is that a user may use a resource or service without being distinguished from other users and without disclosing his identity to third parties. In order for anonymity to be achieved, data communicated in the sensor network needs either to be depersonalised or pseudonymised, in cases that the legitimate network service requires user identification.

D. User's notice and choice

In order for a sensor network deployment to be trusted, it needs to provide awareness to the individuals within it that data is being collected for them or the environment surrounding them. It also has to empower them to control, by making informed decisions, what personal data will be disclosed and which of the network pseudonyms representing them should be used. From the user's side, this could be achieved through possessing a privacy enhancing identity management system [3] that should be more sophisticated than traditional sensor nodes both in terms of computational power and user interaction capabilities, that would act as a gateway between his body sensor network and the environment the sensed data is released to. Trust in a sensor network would be enhanced further if it made its privacy policies available and even certified by a trusted third party agency, in order to inform the user about the service provider's the data acquisition and handling practices.

E. Protection of the network's services

An attacker aiming to make a legitimate service violate the user's private space could do this by disrupting its operation and triggering incorrect actions either through masquerading and inserting false data or by altering the measurements sent by legitimate nodes. In order to protect the user's privacy, it is thus crucial to protect the network services from malicious

intrusions. What mainly has to be ensured is the authenticity, integrity and freshness of the data collected, communicated or calculated by aggregator nodes.

III. CHALLENGES IN PRESERVING PRIVACY

Several security mechanisms, tools and applications have been proposed to enhance Web privacy and anonymity [4]. For sensor networks, however, these solutions can not be directly applied. Their inherent properties in node, network and data level pose challenges that are unique in the networks security area. The following issues are the ones usually encountered when designing security architectures that aim to protect, among others, the privacy of the network's users:

A. Sensor node capabilities

Sensor nodes are designed to be small and inexpensive and are thus constrained regarding their energy, memory, computation and communication capabilities. This poses limitations on the range of cryptographic primitives they can support. Traditional public key cryptography, while being well suited to fulfill requirement (A), is considered unrealistic for sensor networks [5]. Techniques like onion routing [6], aiming to protect the sender's identity (requirements (B) and (C)), would impose high computational overhead to all network components. The use of dummy traffic [7], that has been proposed in order to achieve indistinguishability (requirement (C)), would exhaust the energy supplies of the sensor nodes. Moreover, since sensor networks are often deployed in accessible areas, the nodes themselves are physically vulnerable. By compromising a sensor node, an attacker could obtain its cryptographic keys or even reprogram it, thus obtaining an authenticated malicious node.

B. Communication issues

The wireless nature of sensor network communications makes it even more challenging to fulfill most of the requirements set. It makes them vulnerable not only to eavesdropping attacks but also to false data injection, since an adversary does not need to gain physical access to the networking infrastructure. Moreover, sensor networks are infrastructure-less and dynamic. The lack of central servers and static base stations, the dynamically changing network topology, the possibility of addition or deletion of sensor nodes through all stages of their life cycle, all combined with the scale of the deployments (hundreds or thousands of sensor nodes), set strict requirements on the authentication and encryption schemes that can be used: they need to be distributed, flexible, scalable, and cooperative. The lack of centralised host relationships also affects the routing mechanisms of the data packets, that will typically follow multi hop routes before arriving at their final destination. The trustworthiness of the intermediate nodes of each path can not be guaranteed a priori.

C. Data handling

In order to minimise communication costs in sensor networks, large streams of data are converted to aggregated

information within the network by data aggregator nodes. The fact that in-network processing is performed sets additional security requirements for node-to-aggregator node communication. In traditional networks, where each node communicates with some base station, most data integrity and identity confidentiality requirements can be fulfilled by end to end encryption, message authentication codes and the use of pseudonyms. In sensor networks, however, for data aggregation purposes, intermediate node authentication and message integrity verification will be required. Issues related to how user pseudonyms can be handled in the presence of (possibly untrusted) aggregator nodes and which raw or aggregated data can be anonymised at each point within the network also need to be resolved.

IV. DESIGNING PRIVACY PRESERVING SENSOR NETWORKS

The unique challenges faced in sensor networks have inspired research on the design of lightweight and flexible security solutions. Towards fulfilling sensor networks privacy requirements, the solutions that have been proposed are identified and categorised to correspond to the requirements discussion in Section II.

A. Ensuring Data Confidentiality

Data encryption is the typical defense against eavesdropping, and the only method for preserving the confidentiality of exchanged messages' content. The basic prerequisite for encryption to be applied is the sharing of cryptographic keys between the communicating parties. The distribution of cryptographic keys to sensor networks is an active research area, since traditional key distribution schemes can not be directly applied. For a solution to be viable given the characteristics and limitations of sensor networks, it needs to entail acceptable computational overhead, which excludes the use of traditional public key cryptography algorithms like RSA or Diffie Hellman key agreement. Moreover, it both has to be scalable and flexible, thus not based on one trusted certification authority, and its memory requirements should be low, proving the sharing a unique symmetric key between each pair of nodes to be impractical. Finally, it has to provide some level of resilience to node compromise, thus excluding the use of a network wide shared key.

The solution that prevails toward meeting the security requirements in sensor networks is symmetric key encryption, due to its low computational cost and memory requirements. One of the most well known security architectures that utilizes symmetric cryptography is SPINS [8], where asymmetry is introduced into symmetric key cryptography through delayed key disclosure and one-way function key chains. A fully distributed key management scheme for sensor networks was introduced in [9]. It relies on probabilistic key predistribution, where each sensor node is assigned a random subset of keys from a key pool before deployment, so that each pair of sensor nodes has a certain probability to share at least one symmetric key. Various improvements and extensions to the

basic probabilistic scheme have been proposed [10][11]. Other approaches to symmetric key establishment include LEAP [12], where keys with different scope are used to provide different security levels, LiSP [13], which includes a rekeying mechanism periodically performed by group-head nodes, and PIKE [14], which facilitates pairwise key establishment using peer sensor nodes as trusted intermediaries.

Although traditional public key cryptography was initially considered inapplicable for sensor networks, Elliptic Curve Cryptographic key generation emerged as an attractive alternative that would allow for greater scalability and flexibility [5], while being efficient enough to be attained and executed on resource-constrained sensor nodes [15], mainly due to the fact that it can offer equivalent security with smaller key sizes. Provided the infrastructureless nature of sensor networks, a challenging concern related to applying public key cryptography is what will comprise the certification authority. One approach is a fully distributed selforganizing public key management system that uses no certification authority at all [16]. Instead, nodes are responsible for generating their keys, and then issuing and distributing their public-key certificates. However, although it is a simple and efficient solution, it can not fulfill strong authentication and accountability requirements. Another approach [17] is the use of a virtual certification authority, distributed using threshold cryptography [18] among a set of cooperating selected nodes, each holding a share of the service private key.

B. Protecting the communication's context

Independently of what encryption scheme is being used, the cipher texts should not allow induction of any information related to their context. Techniques like timestamping, padding, using serial numbers, or frequent key redistribution can be used so that the communicated cipher texts do not reveal information through of their similarity or size. Protecting the traffic patterns within the network, which includes the network identities of the sender and recipient of data packets, is not trivial, as it requires interference with the routing protocol.

One of the earliest techniques proposed for untraceable communication through electronic mail is MIXEs [19]. It is based on independent servers that, after collecting a number of messages, mix them by applying cryptographic operations to change their outlook and then forward them in a different order. However, for this technique to be applied in sensor networks, it would have to be ensured that every sensor is within the range of such a server, which would still not protect from eavesdropping on the path between them.

A secure routing protocol that uses the onion routing technique to protect the identity of both the sender and the recipient from the intermediate routing nodes is presented in [20]. Assymmetric encryption is used by the sender in an onion-like form for the network path to the recipient, in a way such that each intermediate node, after decrypting the packet received using its private key, can only identify the next hop in the path without being able to determine whether the next hop is the final destination. The basic concern related to applying this

technique in sensor networks is the computational overhead it would impose to all nodes in a network path.

Furthermore, a privacy preserving network needs to protect not only the network identities of the users, but also the information related to their actual or relative locations. Even legitimate location tracking systems, used to provide location-based services, may either be passively or actively attacked or misused. For services that only require location statistics without user identification, eg traffic control, user's location data can be anonymised [21], in which case their location privacy depends on the number of people around them, as they comprise their anonymity set. For services that require user identification, trusted user agents can be used as intermediaries, applying user-defined policies on location based requests [22]. Although the definition of user policies by the users may prove to be unrealistically complex, such an approach would give the user primary control over its location information.

C. User's anonymity and pseudonymity

In the context of sensor networks, anonymity mechanisms allow the individuals to use the network services through the nodes that are related to them, while protecting their identity from possible abuse. Restricting the network's ability to gather data at a detail level that could compromise it through depersonalising the results reported by sensors [23], is a first step toward fulfilling this requirement. For legitimate network services that require user identification, the identity of the user can be secretly shared between him and the service provider through the use of pseudonyms.

The scope of pseudonyms may differ according to the context for their use, providing different strength of long-term linkability [2]. A static person pseudonym would be the weakest, while the establishment of transaction pseudonyms can be used to achieve strong unlinkability. A role-relationship pseudonym, different for each service provider and user role is an intermediate case that would protect against identity information correlation from the service provider's side. Pseudonyms may be used to represent users of groups of users, eg households, in which case the group also represents each user's anonymity set. The establishment of pseudonyms and the binding to their holders can be performed by a special kind of certification authority, for the deployment of which in sensor networks the concerns and solutions described in Section IV-A apply.

However, even if data is anonymised or pseudonyms are being used, each user's anonymity depends on his anonymity set: if it shrinks, his identity can be disclosed. A worst case scenario is for a user to be alone in the an area of a sensor network. All messages sent by his nodes will be directly linked to him and the pseudonym he uses will be disclosed. A solution would be to use, together with the anonymity mechanisms, dummy traffic or background noise [7]. However, for the resource constrained sensor nodes, the use of dummy traffic would be unsuitable. If the disclosure of pseudonyms can not be avoided when the anonymity set shrinks, the

periodical or user-requested pseudonym redistribution might be the only solution.

D. Allowing for user-controlled data disclosure

Empowering sensor network users to control the level of privacy according to the context, their role and communication partner mainly entails two actions. Firstly, mechanisms should be provided to inform them whether data collection is being performed and what privacy policies are being announced. Having provided the users with access to information about their privacy risks, mechanisms that would allow the definition and application of their preferences would enable them to control if any of their personal data should be disclosed, using which one of their pseudonyms.

The provision of such functionality requires the addition of extra components in the general case of sensor networks architecture: the privacy assistants, acting as the user gateways to the various sensor network applications. PDAs are proposed to be used as gateways in [24], as part of a scheme to provide awareness of possible privacy threats. The scheme, based on the notion of secure two-party point-inclusion problem, enables the user to conclude whether he is inside the sensing areas of some type of sensor networks, without disclosing his exact position within the area. A scheme that allows for user control, assuming that the sensor network can be trusted and cooperating, is proposed in [25]. It includes mechanisms for the network service to announce its privacy policies and data handling practices, and for the users to apply their privacy preferences on accepting or declining a service. It is, however, set as a prerequisite that the services are optional and configured to suit the users' decisions related to their privacy. Moreover, the mechanism that would allow users to configure their preferences in an easy, understandable manner has not been described yet.

E. Ensuring data authenticity and integrity

Data authenticity mechanisms typically depend on the key distribution scheme (Section IV-A), while data integrity is ensured by the use of hash functions. In sensor networks, the possibility of node compromise poses additional threats, since the compromised nodes can be authenticated to the network. For example, a stelthy attack, where an attacker's goal is to make the network accept a false data value using a compromised node, combined with a sybil attack, where a malicious node illegitimately claims multiple identities, would allow one compromised node to have a greater impact on causing a false aggregation result [26].

At the same time, sensor networks allow for redundancy on the views of the environment, which can be exploited for ensuring data correctness either by using majority voting between the nodes that were around a reported event, or by cross-checking the collected results for consistency. This characteristic is exploited in a scheme proposed in [27] in order to ensure the validity of the results provided by aggregator nodes, where the MACs of witness nodes that conduct the same data fusion operations as the aggregators are used as proofs.

A reactive scheme aiming to ensure the results provided by aggregator nodes are good approximations of the true values, even if the aggregators and a fraction of the sensor nodes are compromised, is proposed in [28]. It is based on representing the data used for the aggregation using Merkle hash trees, that the base station can verify using random sampling mechanisms and interactive proofs. The problem of compromised nodes tampering with the transmitted data on the network path was studied in [29], where an authentication scheme is proposed to guarantee that the base station will detect any injected false data packets when no more than a certain number of nodes are compromised.

V. CONCLUDING REMARKS

From the schemes that have been proposed, it becomes apparent that there exist solutions to fulfill most of the requirements set. However, some issues can not be disregarded; Firstly, most of the schemes presented, especially for protecting the context of the communications, influence the system design at the networking protocol level, which complicates their actual integration to the deployments. Moreover, issues related to anonymising or pseudonymising data depend on the application domain, the in-network data processing schemes and the privacy sensitivity of each user. Thus, it may be infeasible to design a generic and high level privacy architecture, that could both be independent of the underlying networking protocols and guarantee some level of privacy independently of the context of the deployment.

Another issue is related to the level of trust users need to have to the deployments in order to take full advantage of the services that can be offered. The definition by the users of strict privacy policies that would guarantee that personal information is not disclosed, would also not allow them to use legitimate services that require that information. It would thus be necessary to build some level of trust to legitimate deployments, which can not be accomplished using solely technical means. Trusted privacy certification authorities, the appropriate legal deterrence and societal norms are expected to help toward this direction in the future.

REFERENCES

- [1] A. Perrig, J. Stankovic, and D. Wagner, "Security in wireless sensor networks," *Communications of the ACM*, vol. 47, no. 6, pp. 53–57, 2004.
- [2] A. Pfitzmann and M. Hansen, "Anonymity, unlinkability, unobservability, pseudonymity, and identity management – A consolidated proposal for terminology," February 2006, version v0.27. [Online]. Available: http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.27.pdf
- [3] M. Hansen, P. Berlich, J. Camenisch, S. Clauß, A. Pfitzmann, and M. Waidner, "Privacy-enhancing identity management," *Information Security Technical Report*, vol. 9, no. 1, pp. 35–44, 2004.
- [4] S. Gritzalis, "Enhancing Web privacy and anonymity in the digital era," *Information Management and Computer Security*, vol. 12, no. 3, pp. 255–287, 2004.
- [5] B. Arazi, I. Elhanany, O. Arazi, and H. Qi, "Revisiting public-key cryptography for wireless sensor networks," *IEEE Computer*, vol. 38, no. 11, pp. 103 – 105, November 2005.
- [6] D. M. Goldschlag, M. G. Reed, and P. F. Syverson, "Hiding routing information," in *Proceedings of the 1st International Workshop on Information Hiding*. Springer-Verlag, 1996, pp. 137–150.
- [7] A. Pfitzmann, B. Pfitzmann, and M. Waidner, "ISDN-MIXES – Untraceable communication with very small bandwidth overhead," in *Proceedings of the 7th IFIP International Conference on Information Security*. Elsevier, 1991, pp. 245–258.
- [8] A. Perrig, R. Szewczyk, J. D. Tygar, V. Wen, and D. E. Culler, "SPINS: Security protocols for sensor networks," *Wireless Networks*, vol. 8, no. 5, pp. 521–534, 2002.
- [9] L. Eschenauer and V. D. Gligor, "A key-management scheme for distributed sensor networks," in *Proceedings of the 9th ACM Conference on Computer and Communications Security*. ACM Press, 2002, pp. 41–47.
- [10] H. Chan, A. Perrig, and D. X. Song, "Random key predistribution schemes for sensor networks," in *Proceedings of the IEEE Symposium on Security and Privacy*, 2003, pp. 197–213.
- [11] D. Liu and P. Ning, "Improving key predistribution with deployment knowledge in static sensor networks," *ACM Transactions on Sensor Networks*, vol. 1, no. 2, pp. 204–239, 2005.
- [12] S. Zhu, S. Setia, and S. Jajodia, "LEAP: Efficient security mechanisms for large-scale distributed sensor networks," in *Proceedings of the 10th ACM Conference on Computer and Communications Security*, 2003, pp. 62–72.
- [13] T. Park and K. G. Shin, "LiSP: A lightweight security protocol for wireless sensor networks," *ACM Transactions on Embedded Computing Systems*, vol. 3, no. 3, pp. 634–660, 2004.
- [14] H. Chan and A. Perrig, "PIKE: Peer intermediaries for key establishment in sensor networks," in *Proceedings of IEEE Infocom*, 2005.
- [15] A. Wander, N. Gura, H. Eberle, V. Gupta, and S. C. Shantz, "Energy analysis of public-key cryptography for wireless sensor networks," in *Proceedings of the 3rd IEEE International Conference on Pervasive Computing and Communications*, 2005, pp. 324–328.
- [16] J.-P. Hubaux, L. Buttyán, and S. Capkun, "The quest for security in mobile ad hoc networks," in *Proceedings of the ACM Symposium on Mobile Ad Hoc Networking and Computing*, 2001, pp. 146–155.
- [17] S. Yi and R. Kravets, "Moca: Mobile certificate authority for wireless ad hoc networks," in *Proceedings of 2nd Annual PKI Research Workshop*, 2003.
- [18] A. Shamir, "How to share a secret," *Communications of the ACM*, vol. 22, no. 11, pp. 612–613, 1979.
- [19] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–88, 1981.
- [20] B. Awerbuch, D. Holmer, C. Nita-Rotaru, and H. Rubens, "An on-demand secure routing protocol resilient to byzantine failures," in *Proceedings of the ACM Workshop on Wireless Security*, 2002.
- [21] M. Gruteser, G. Schelle, A. Jain, R. Han, and D. Grunwald, "Privacy-aware location sensor networks," in *Proceedings of the 9th Workshop on Hot Topics in Operating Systems*, 2003, pp. 163–168.
- [22] M. Spreitzer and M. Theimer, "Providing location information in a ubiquitous computing environment (panel session)," in *Proceedings of the 14th ACM Symposium on Operating Systems Principles*. ACM Press, 1993, pp. 270–283.
- [23] H. Chan and A. Perrig, "Security and privacy in sensor networks," *IEEE Computer*, vol. 36, no. 10, pp. 103–105, 2003.
- [24] Y. Sang and H. Shen, "A scheme for testing privacy state in pervasive sensor networks," in *Proceedings of the 19th International Conference on Advanced Information Networking and Applications*. IEEE Computer Society, 2005, pp. 644–648.
- [25] M. Langheinrich, "A privacy awareness system for ubiquitous computing environments," in *Proceedings of the Ubiquitous Computing International Conference*, 2002, pp. 237–245.
- [26] E. Shi and A. Perrig, "Designing secure sensor networks," *Wireless Communication Magazine*, vol. 11, no. 6, December 2004.
- [27] W. Du, J. Deng, Y. S. Han, and P. K. Varshney, "A witness-based approach for data fusion assurance in wireless sensor networks," in *Proceedings of the IEEE Global Telecommunications Conference*, vol. 3, 2003, pp. 1435–9.
- [28] B. Przydatek, D. Song, and A. Perrig, "SIA: secure information aggregation in sensor networks," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems*. ACM Press, 2003, pp. 255–265.
- [29] S. Zhu, S. Setia, S. Jajodia, and P. Ning, "An interleaved hop-by-hop authentication scheme for filtering of injected false data in sensor networks," in *Proceedings of the IEEE Symposium on Security and Privacy*, 2004, pp. 259–271.