On Providing Anonymity in Wireless Sensor Networks

A. Wadaa[1], S. Olariu[2], L. Wilson[3], M. Eltoweissy[4], K. Jones[5]

{1,2,3} Department of Computer Science, Old Dominion University, Norfolk, VA 23529
{4} Department of Computer Science, Virginia Tech, Falls Church, VA 22043
{5} NASA Langley Research Center, Hampton, VA 23681
wadaa@cs.odu.edu

Abstract

Securing wireless sensor networks against denial of service attacks that disrupt communications or target nodes serving key roles in the network, e.g. sinks or routers, is instrumental to network availability and performance. Particularly vulnerable to these attacks are the components of any communications or operation infrastructure in the network. In this paper, we address a class of wireless sensor networks where network protocols leverage a dynamic general-purpose virtual infrastructure; the core components of that infrastructure are a coordinate system, a cluster structure, and a routing structure. Since knowledge of this virtual infrastructure enables 'smart' cost-effective DOS attacks on the network, maintaining the anonymity of the virtual infrastructure is a primary security concern. The main contribution of this work is to propose an energy-efficient protocol for maintaining the anonymity of the network virtual infrastructure. Specifically, our solution defines schemes for randomizing communications such that the coordinate system, cluster structure, and routing structure remain invisible to an external observer of network traffic during the setup phase of the network.

1. Introduction

Recent advances in nano-technology made it technologically feasible and economically viable to develop low-power devices that integrate general-purpose computing with multiple sensing and wireless communications capabilities. Each of these devices, called a sensor node, packs a limited non-renewable power supply and once deployed, must work unattended. We envision a massive random deployment of these commodity sensor nodes, numbering in the thousands or tens of thousands. Aggregating this large number of sensor nodes into sophisticated computation and communication infrastructures, called sensor networks, will have a significant impact on a wide array of applications including military, scientific, industrial, and health. The fundamental goal of a wireless sensor network is to produce, over an extended period of time, meaningful global information from local data obtained by individual sensor nodes [1,3,7,10,14].

However, a wireless sensor network is only as good as the information it produces. In this respect, perhaps the most important concern is information security. Indeed, in most application domains sensor networks will constitute a mission critical component requiring commensurate security protection. Information security in wireless sensor networks is exacerbated by a number of factors: (1) Sensor nodes are resource constrained in terms of computation, communication, and energy, (2) sensor nodes are highly vulnerable to physical tampering, (3) sensor nodes communicate through insecure wireless links, and (4) being commodity devices, individual nodes may not even have unique identifiers, and immediately after deployment are unaware of their location [1,3,9,12,21]. An information security solution for wireless sensor networks must protect against an adversary perturbing the information produced, stopping production, or pilfering information, taking into account all these factors.

In many sensor network applications, safeguarding output data assets, i.e. data produced by the sensor network and consumed by the end user (application), against loss or corruption is the main security concern. In these applications, a sensor network is typically deployed into a hostile target environment for an amount of time. The network self-organizes and works to generate and forward, or store for later access, output data of import to the application. As an example, a sensor network may be deployed across enemy territory ahead of a planned attack; the network system monitors the environment and produces and stores reconnaissance data that is material to mission planning. Periodically, during the network lifetime, a mobile gateway, mounted on a person, land or

IEEE COMPUTER SOCIETY

airborne vehicle, or a satellite, may collect the output data assets from the network system to maintain real-time situational awareness. In this scenario the network system must store the output data assets from the time it is produced until it is collected. Therefore, securing the output data assets in the network is an important problem in this type of applications [1,3,16].

We model an attack on the output data assets in the sensor network as a type of denial of service attacks. This is based on the abstraction that output data is stored in a logical repository, and, that access to this output data repository constitutes, in effect, a service provided by the network system to the application; corruption or loss of output data denies the application access to that service. Many wireless sensor networks are mission-oriented, must work unattended, and espouse data-centric processing. Consequently, they are significantly different in their characteristics from conventional ad-hoc networks. Security solutions designed specifically for wireless sensor networks are therefore required.

1.1 What is anonymity?

In a communications system, *anonymity* typically refers to maintaining the identity of one or more parties in a communication anonymous to other parties involved or not involved in that communication. Recent years have seen a flurry of activity and many communication systems that maintain some type of anonymity have been developed for the Internet [4-6,8,11,13,15,17-19]. Most of the work on anonymity can be broadly classified into *sender* anonymity, *receiver* anonymity, or *mutual* anonymity. In *e-voting*, and Internet-based access to public information, sender anonymity is a primary concern. In private transaction-based applications, e.g. banking, keeping both the sender and receiver anonymous to a third party is necessary. The rapid growth of Internet-based connectivity and applications projected *traffic* anonymity as an important concern. If an adversary can identify traffic (or traffic patterns) associated with particular applications, then it can easily leverage this knowledge to jeopardize the anonymity of communicating parties, mount targeted security attacks on either the traffic, or the network infrastructure components supporting that traffic, e.g. routers, bridges, etc.

Recently, the problem of securing ad-hoc networks has received a great deal of well-deserved attention in the literature [1,3,9,12,21]. Somewhat surprisingly, however, in spite of its importance, anonymity problems have not been addressed specifically in wireless sensor networks.

We view this work as an initial contribution towards developing an efficient solution for the anonymity problem in wireless sensor networks. In this paper, we focus primarily on *structure* anonymity in a wireless sensor network. Structure anonymity refers to the anonymity of any communications or operations infrastructure in the network to an external observer. It should be noted that we assume a wireless sensor network where a *dynamic virtual infrastructure*, overlaid on top of the physical sensor nodes, is leveraged by network protocols. Although the exact statement of the problem we address is presented in Section 5, suffice it to say that the basic elements of such virtual infrastructure are a coordinate system that affords natural clustering, and a routing structure. Our main contribution in this paper is an efficient scheme to maintain the anonymity of the coordinate system, and cluster and routing structures during the network setup phase, in the presence of an external adversary. Our solution accomplishes this by securing the process that the sensor nodes use to acquire knowledge of the virtual infrastructure, and minimizing communications involving the sensor nodes. In fact, under our solution, the acquisition process does not involve *any* transmission of messages from sensor nodes!

2. The network model

The network model used in this paper is based on the model introduced in [20]. Specifically, we assume a class of wireless sensor networks consisting of a large number of sensors nodes randomly deployed in the environment of interest. A training process, as explained below, establishes a coordinate system and defines a clustering of all nodes. Post training, the network undergoes multiple operation cycles during its lifetime. The training process also endows the role of *sink* upon one or more of the defined clusters. The sink role is transient, however, since new sink clusters are designated at the beginning of each operation cycle. Each sink cluster, henceforth called sink, acts as a repository for a portion of the sensory data, generated in the network during an operation cycle. At the end of an operation cycle, each sink offloads data stored in its repository to a *gateway*. In the following we describe the three primary entities in our network model in more detail.

2.1 The sensor node

We assume that individual sensor nodes have four fundamental constraints: (1) sensors are anonymous; initially a sensor node has no unique identifier, (2) each sensor has a limited non-renewable energy budget, (3) each sensor attempts to maximize the time it is in sleep mode; a sensor wakes up at specific (possibly random) points in time for short intervals under the control of a timer, and (4) each sensor has a modest transmission range, perhaps a few meters with the ability to send and receive over a wide range of frequencies. In particular, communication among sensor nodes in the sensor network must be multi-hop.

2.2 The sink

In our network model, all nodes in a sink cluster serve as sink nodes. Coarser sink granularity means higher sink storage capacity, and potentially longer operation cycles. However, coarser sink granularity, as envisioned here, comes at a potentially higher risk of anonymity attacks due to the space correlation among sink nodes. If it is discovered that a node, x, is a sink node then it immediately follows that there exists at least one other sink node (typically more) in the vicinity of x. Thus, the success of an anonymity attack in our model is a function of the probability of identifying the *first* node in a sink. One approach to decrease the latter probability is to have only a subset of the nodes in a sink cluster serve as sink nodes. There is a tradeoff between sink granularity, and the amount of security a sink has against anonymity attacks.

2.3 The gateway

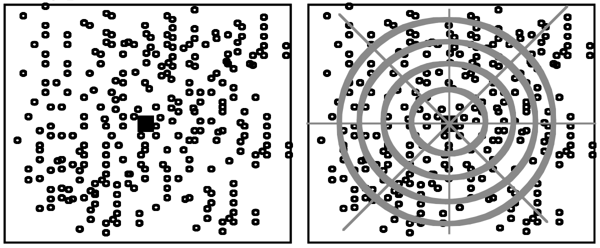
The gateway is an entity that connects the sensor network system to the outside world. The gateway is not constrained in mobility, energy, computation, or communication capabilities. There are two basic functions for the gateway in our network model:

i. *Training*: The gateway performs the network training process, post deployment. For training purposes, under our solution, the gateway does not need to be within transmission range of the nodes, and is assumed to be able to send long-range, possibly, directional broadcasts to all sensors,

ii. *Harvesting* (data collection): At the end of an operation cycle, the gateway typically collects sensory data stored in each sink. In a simple collection scenario the gateway traverses the deployment environment to collect data from all the sinks.

3. Network organization and clustering

Figure 1(a) features an untrained sensor network immediately after deployment in an environment that is modeled here by a 2-dimensional plane. With no loss of generality, we assume that the trainer is centrally located relative to all deployed nodes. The primary goal of training is to establish a *coordinate system*, to provide the nodes location *awareness* in that system, and to organize the nodes into *clusters*. The coordinate system, and clustering are briefly explained next. We refer the interested reader to [20] for an in-depth description of the training process.



(a) **(b)**

Figure 1: (a) an untrained sensor network, (b) a trained sensor network

3.1 The coordinate system

The training process establishes a polar coordinate system as exemplified by Figure1 (b). The coordinate system divides the sensor network area into equiangular *wedges*. In turn, these wedges are divided into sectors by means of concentric circles or *coronas* centered at the trainer location. Corona radii can be determined based on several criteria, e.g. in [20] they are designed to maximize the efficiency of *sensors-to-sink* multi-hop communication. The intersection of every wedge and corona defines a unique sector; each sector is uniquely identifiable by the combination of its unique wedge identifier, and unique corona identifier. The training process guarantees that each node belongs to one and only one sector in the coordinate system, and that each node knows the identity of its sector [20].

Let *c,* and *w* be, respectively, the set of coronas, and the set of wedges defined by the training process. The resulting coordinate system can thus be formally represented by $\{(r_0, r_1, \ldots, r_{|c| \ 1}), \ \}$, where r_i is the radius of corona *i,* $0 \quad i \quad |c| \quad 1$, is the wedge angle, and $|w| = 2□/□$. A fundamental assumption here is that any coordinate system is designed such that all nodes located in the same sector can communicate using direct (single hop) transmission.

3.2 Clustering

A major advantage of our coordinate system is that sectors implement the concept of clustering (at no additional cost). A sector effectively constitutes a cluster;

clusters are disjoint, and are uniquely identifiable. All nodes located in the same sector are members of the same cluster, and have the same location coordinates, namely, the corona and wedge identifiers corresponding to that sector. This clustering scheme is ideally suited for sensor nodes that are intrinsically anonymous. We proposed in [20] a scalable training protocol where each untrained node incurs a communication cost equal to *log*|*w*|+*log*|*c*|, and the nodes do not transmit any messages during the training process.

4. The work model

The work model defines how sensor nodes work collaboratively to generate and store sensory data during an operation cycle. We propose a model that divides work into i*ntra-cluster activity (data generation)*, and *inter-cluster activity* (data transportation for in-network storage)

4.1 Intra-cluster activity

In our model, the sensory data resulting from intra-cluster activity encodes states of a process of interest. Namely, we assume that the goal of intra-cluster activity is to monitor a process (or phenomenon), and report on its *local state* at any point in time. The state space of the phenomenon is given by $\{s_0, s_1, s_2, ..., s_z\}$. State s_0 denotes the *normal state*, and each of s_i, $1 \quad i \quad z$, denotes an *exception state.* The assumption here is each state s_i, $1 \square i \square z$, corresponds to an application-defined exception of a particular type. The normal state corresponds to the fact that no exception of any type is detected.

We propose a transaction-based model for managing the computation and reporting of target process states. The model is a specialization of a transaction-based management model for sensor networks introduced in [20]**.** In this model, intra-cluster activity proceeds as follows. For a given cluster, subsets of nodes located in the cluster dynamically band together forming *workforces*. Periodically, members of each workforce collaborate to perform an instance of a state computation transaction preloaded into each node. The transaction computes and reports the local process state. Note that the system allows for a fresh transaction to be downloaded to the nodes at the beginning of each operation cycle. In the simplest case, performing an instance of the state computation transaction entails that each member in the corresponding workforce perform a sensing operation and formulate a node report. A specific member of the workforce, designated as a transaction instance manager, then receives all node reports, and formulates a *Transaction Instance Report (TIR)*. The TIR is the encoding of the local process state of interest at the time. This TIR is subsequently transported to a sink for storage. In principle, after transmitting the TIR the corresponding workforce disbands. For simplicity, we assume that at most one transaction instance is in progress in a given cluster at any given point in time.

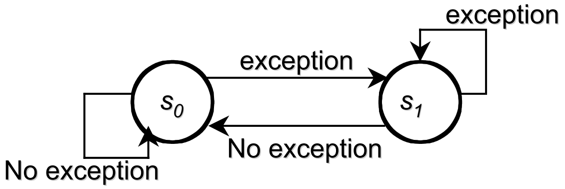


Figure 2. example workforce behavior (assumes a two-state target process, and a canonical transaction , T)

The state diagram in Figure 2 illustrates the behavior of an arbitrary workforce in a cluster assuming the special case of a target process that has one normal state, s_0, and one exception state, s_1. T in the figure denotes the state computation transaction. A more detailed discussion of the design issues pertaining to this work model, including negotiating workforce size subject to QoS constraints, and workforce setup schemes can be found in [20].

4.2 Inter-cluster activity

As indicated earlier, the goal of inter-cluster activity in our work model is to route TIRs from their clusters of origin to the sinks, by means of multi-hop communication. We define a hop in a route as a direct transmission from one cluster to a *neighbor* cluster. A cluster u_j is a neighbor of a cluster u_i if and only if both u_j and u_i are located in the same corona, or the same wedge, for all $i \text{ and } j, i \quad j$. It follows that in any instance of the coordinate system defined in Section 3, each cluster has either three or four neighbors. The set of all neighbors of a cluster u_i is called the *neighborhood* of u_i. In the our proposed anonymity solution we define a distributed inter cluster routing protocol that yields optimal routes in terms of the number of hops from source to sink, and is highly scalable in the number of clusters in the network. Scalability can be attributed to two characteristics of the protocol. First, the protocol uses *no* dynamic global or regional state information. This eliminates the need for control messages to support routing. Second, the protocol uses distributed (incremental) route computation; the destination of hop i, computes, in turn, the destination of hop $i+1$, $i \quad 1$. In the remainder of this paper we assume the availability of a MAC layer that supports inter-cluster communication.

5. The anonymity problem

In this section we formulate a definition for the anonymity problem addressed in this paper. The problem is defined in the context of the network system described earlier. First, we introduce an anonymity threat model.

5.1 The anonymity threat model

The threat model assumed in this paper emanates from a data-centric view of the sensor network. The model is predicated on the assumption that *the end-goal of anonymity attacks on the sensor network is to identify and eliminate the minimum number of nodes to inflict maximum loss of data assets*; In our sensor network model, TIRs are the data assets of import to the end user (application). For any operation cycle, if a sink suffers a permanent failure before transferring the contents of its TIR repository to the gateway, then a portion of the data assets corresponding to the cycle is irrevocably lost. Therefore, eliminating sinks, or nodes comprising them to be precise, is the end-goal of anonymity attacks in our model. There are two main approaches to eliminating sink nodes, as follows.

Brute-Force (Sink nodes are not identified): This may take the form of randomly eliminating nodes in the network on the assumption that, statistically, some sink nodes will be eliminated in the process. Coarse sink granularity, and sink redundancy mitigate the risk of loss of data assets as a result of this type of attack.

Smart (Sink nodes are identified): The adversary system analyzes network traffic to discover, i.e. compromise the anonymity of, sinks, and, hence, eliminate them.

In this paper we assume the adversary system engages in smart elimination attacks.

5.2 Terminology and notation

The main goal of this subsection is to establish terminology and notation that will be used in our solution to the anonymity problem.

$Trudy$ Denotes the adversary system

m A message transmitted in the sensor network system post deployment. (Note that the transmitter is either a sensor node or the gateway.)

$ts(m)$ A global time stamp assigned by Trudy to message m

$l(m)$ The unique location in Trudy's coordinate system of the transmitter of message m in the 2-dimensional deployment plane.

$r^{node}, r^{gateway}$ The nominal transmission radius of a sensor node, and the gateway, respectively

$cvr(m)$ This is the *cover* of message m. Specifically, $cvr(m)$ is the set of nodes that are located in the circular area with radius r^{node}, and center $l(m)$ in the deployment plane

$trace(m)$ The trace of message m; if m is routed along a path of length g, $trace(m)$ is the sequence of messages $\left(m^{(1)}, m^{(2)}, \cdots, m^{(g)}\right)$, $m^{(i)}$ is the retransmission of m, or an encryption thereof, over hop number i, $1 \quad i \quad g$. Note that $m^{(1)} = m$

$source(m)$ The transmitter of message m

$destination(m)$ The intendeds receiver of message m

The assumptions underlying our anonymity threat model can be summarized as follows:

Pre-deployment:

i. All nodes are trusted,
ii. Nodes are in a secure environment,
iii. Trudy does not have access to any message transmitted in the system,

Post deployment (training and operation cycles):

iv. Trudy receives every message transmitted in the system. Note that receiving a message does not imply being able to interpret it. A message m transmitted in the system is represented in Trudy's system as follows: $\left(m, ts(m), l(m), cvr(m)\right)$

5.3 Anonymity problem statement

Let I be an arbitrary time interval, that starts post deployment, and let the set $M = \left\{\left(m_i, ts(m_i), l(m_i), cvr(m_i)\right) \mid 1 \quad i \quad h\right\}$ be the set of all messages transmitted in the system (and hence recorded by Trudy) during the interval I. Also, let the coordinate system, O, established by training be given by $\{(r_0, r_1, \ldots, r_{|c|\square 1}), \square\}$, and the set of all nodes located in sink clusters be denoted by S. The anonymity problem (from Trudy's point of view) can thus be stated as follows.

Given:

$$M = \{(m_i, ts(m_i), l(m_i), cvr(m_i)) | 1 \quad i \quad h\} \qquad (1)$$

Find:

$$m_q \quad M:$$

$$(\; m_p \quad M:$$

$$trace(m_p) = (m_p^{(1)}, m_p^{(2)}, \cdots, m_p^{(g_p)}) \quad m_q = m_p^{(g_p)})$$

Note that for the message m_q in (2), $destination(m_q) \quad S$. In general, for each message m_q that satisfies (2), it follows that

$$x : x \quad cvr(m_q) \quad x \quad S$$

The challenge for the sensor network system is to devise training, intra cluster, and inter cluster protocols that minimize the probability that the anonymity problem stated in equations (1), and (2) is solved, for arbitrary O, I, S, and M. In this work we only look at the problem of providing anonymity during the training period.

6. Providing network training anonymity

The main goal of this section is to propose a protocol for training that addresses the anonymity problem formulated above. The primary goal of the training protocol is to establish the canonical coordinate system, O_s, for the network, anonymous to Trudy. For a given sensor network system, the canonical coordinate system is the instance of the polar coordinate system described in 3.4.1 that has the maximum precision. O_s defines the set of canonical coronas, c_s, and the set of canonical wedges, w_s; we assume that $|c_s|$ and $|w_s|$ are powers of 2. O_s is defined by $\{(r_0, r_1, \ldots, r_{|c_s|\square 1}), \square_s\}$, $r_i = (i+1) \quad _s, 0 \quad i \quad |c_s| \quad 1$, where $_s$, and $_s$ are, respectively, the smallest corona width, and the smallest wedge angle for the system; $_s$, and $_s$ characterize the system precision.

Post training, the coordinate system used during any operation cycle is derived from the canonical coordinate system using three integer parameters, $\; , \; , \; and \;$. Here, $\square, \square, and \; \square$ represent, respectively, wedge rotation, wedge grouping, and corona grouping parameters. Let $C(x, O), and W(x, O)$ denote, respectively, the corona, and the wedge where node x is located according to coordinate system O. For a given operation cycle, e, if the rotation and grouping parameters are $\square_e, \square_e, and \; \square_e$, then the coordinate system used for cycle e is defined as follows:

$$O_e = \{(r_0, r_1, \ldots, r_{|c_e| \quad 1}), \quad _e\}, where$$

$$r_i = (i+1) \quad _e, 0 \quad i \quad |c_e| \quad 1, \quad _e = \quad _e \quad _s, \quad _e = \quad _e \quad _s$$

Note that $_e \quad [0, |w_s| \quad 1]$, $_e \quad [0, \log|w_s|]$, and $_e \quad [0, \log|c_s|]$. The corona, and wedge of x according to O_e are defined as follows,

$$C(x, O_e) = C(x, O_s) \; div \; 2^{(\log|c_s| \qquad _e)} \qquad (3)$$

$$W(x, O_e) =$$

$$((W(x, O_s) + \quad _e) \mathrm{mod} \, |w_s|) \, div \; 2^{(\log|w_s| \qquad _e)} \qquad (4)$$

In equation (3) above, $_e$ determines the corona precision of the coordinate system O_e, the minimum precision (a single corona) corresponds to $\square_e = 0$, and the maximum precision (that of the canonical coordinate system) corresponds to $_e = \log|c_s|$. In equation (4) $_e$ determines the wedge precision of O_e in an analogous manner. The left hand operand of the *div* operator in (4) is a translation for $W(x, O_s)$ about the true anchor point an amount equal to $_e$ canonical wedges.

We are now in a position to present the details of our proposed anonymity-compliant training protocol.

Preconditions

i. Sensor nodes are randomly and uniformly deployed in a *deployment area*. The deployment area completely contains a circular area called the *network area*; nodes located in the network area will comprise our trained sensor network. The mission of the nodes that are located outside the network area is to generate fake message traffic to help keep the network area anonymous.

ii. The gateway is mobile

iii. Pre deployment, the following is loaded into each sensor node:

a) Secret key, k_{master} used for decrypting training protocol messages from the gateway. Also, k_{master} is used to derive the keys $k_{\square}, k_{\square}, and\ k_{\square}$, as proposed in [12]. $k\ \ , k\ \ , and\ k$ are used to generate at random values for $\square, \square, and\ \square$, respectively, for the successive operation cycles.

b) The parameters $\ _s, \ _s, D, and\ r^{gateway}$. D is the diameter of the network area, we assume $r^{gateway} >> D$.

iv. Internal clocks in all sensor nodes are synchronized to the gateway.

Training protocol (gateway side)

i. Do a random traversal of the deployment area, visiting a set of random anchor points $A = \{a_1, a_2, \cdots, a_A\}$, For each $a_i, 1 \square i \square | A |, do:$

i.1. *Transmit a call-for-training message*: Transmit an omni directional broadcast message using $r^{gateway}$ The message is encrypted by k_{master} and contains a Boolean flag f that identifies a_i as either the true anchor point or a false anchor point; the true anchor point is the geographical center of the circular network area.

i.2. *Corona train:* using the sink side of the training protocol described in [20] do corona training to establish coronas for O_s, such that

$$| c_s | = \frac{r^{gateway}}{\ _s}$$

i.3. *Wedge train:* using the analogous sink side of the wedge training protocol described in [20], do wedge training to establish the wedges for O_s.

ii. Terminate the protocol

Note that corona training done is step i.2 covers an area considerably larger than the network area. This means that corona training, in this case, defines fake coronas that lie outside the network area (i.e. at a distance more than the diameter *D* from the true anchor point). The objective is to help keep the diameter *D* unknown. Because sensor nodes know *D*, each node, after learning the canonical corona it is located in, can determine if it is located in the network area, and hence, in the trained network. The multiplicity and randomness of anchor points help keep the true anchor point unknown.

Training protocol (node side)

In the following assume node x is the node executing the protocol.

i. *Compute* $| c_s | = r^{gateway} / \ \ _s, | w_s | = 2\ \ / \ \ _s$

ii. *Do forever*:

ii.1. *Receive the next call-for-training message*: receive and decrypt, using k_{master}, the next call-for-training message, m.

ii.2. *If the Boolean flag f in* m *is true*, then do:

ii.2.1. *Get corona trained*: invoke the node side of the training protocol described in [20] to get corona trained to learn the canonical corona number you are located in, $C(x, O_s)$.

ii.2.2. *Compute your corona radius:* Compute $r = (C(x, O_s) + 1)\ \ \ _s$. (Note that if $r \square D$ then you know you are located in the network area, and thus will be a node in the trained network, otherwise you know you do not belong to the trained network.)

ii.2.3. *If you belong to the trained network*, do:

ii.2.3.1. *Compute* $| c_s | = D / \ \ _s$

ii.2.3.2. *Get wedge trained:* invoke the node side of the training protocol described in [20] to get wedge trained to learn the canonical wedge number you are located in, $W(x, O_s)$.

ii.2.3.3. Using $k\ \ , k\ \ , and\ k$ generate via a random number generator (or a preloaded CBC block as described in [12]), respectively, the parameters $\ _1, \ _1, and, \ _1$ for the first operation cycle.

ii.2.3.4. *Compute*

$$C(x, O_1) = C(x, O_s)\ div\ 2^{(\log|c_s| \qquad _1)},$$

$$W(x, O_1) =$$

$$\left(\left(W(x, O_s) + \ \ _1\right) \mathrm{mod}\ | w_s |\right) div\ 2^{(\log|w_s| \quad _1)}$$

ii.2.3.5. *Terminate the protocol.*

```
        Else (do not belong to trained network)
                Sleep for $|w_s|$ message times; terminate
                the protocol.
    Else(flag f in $m$ is false)
            Sleep for $|c_s| + |w_s|$ message times.
```

7. Conclusions

We introduced, and formally defined the problem of structure anonymity in wireless sensor networks where the network protocol leverages a dynamic virtual infrastructure constructed on top of the physical sensor nodes. We proposed an efficient solution for addressing structure anonymity in this class of sensor networks during the network setup phase. Specifically, we developed protocols for ensuring the anonymity of the components of the virtual infrastructure during the time when this infrastructure is being established, and the nodes are acquiring knowledge of that infrastructure. A notable advantage of our solution is that sensor nodes remain completely 'silent' during network setup and virtual infrastructure establishment. Thus our solution is energy-efficient, and scales well in the number of nodes. Currently we are extending our solution to address structure anonymity during the network operation phase.

Acknowledgements: This work was supported in part by a grant from the Commonwealth of Virginia Technology Research Fund (SE 2001-01) through the Commonwealth Information Security Center.

8. References

1. F. Akyildiz, W. Su, Y. Sankarasubramanian, and E. Cayirci, Wireless sensor networks: A survey, *Computer Networks*, 38(4), 2002, 393-422.
2. O. Berthold, A. Pfitzmann, and R. Standke, The disadvantages of free mix routes and how to overcome them, *Proc. Workshop on Design Issues in Anonymity and Unobservability*, July 2000, 27–42.
3. D. W. Carman, P. S. Kruus, and B. J. Matt, Constraints and approaches for distributed sensor network security, TR #00-010, NAI Labs, 2000.
4. E. Gabber, P. B. Gibbons, Y. Matias, and A. Mayer, How to make personalized web browsing simple, secure, and anonymous, *Proc. of Financial Cryptography,* September 1997*,* LNCS 1318, Springer, 1997.
5. D. Goldschlag, M. Reed, and P. Syverson, Onion routing for anonymous and private Internet connections, *Communications of the ACM*, 42, 1999, 39-41.
6. Y. Guan, X. Fu, R. Bettati, and W. Zhao, An optimal strategy for anonymous communications, Dept. of Computer Science, Texas A&M University, Technical Report TR2002-3-1, November 2001.
7. U. Hansmann, L. Merk, M. Nicklous, and T. Strober, *Foundations of Pervasive Computing*, Springer-Verlag, Berlin Heidelbelg, 2002.
8. A. Jerichow, J. Mueller, A. Pfitzmann, B. Pfitzmann, and M. Waidner, Real-time mixes: A bandwidth-efficient anonymity protocol, *IEEE Journal on Selected Areas in Communications,* 16(4), 1998, 495- 509.
9. K. Jones, A. Wadaa, S. Olariu, L. Wilson, and M. Eltoweissy, Towards a new paradigm for securing wireless sensor networks, *Proc. New Security Paradigms Workshop* (NSPW'2003), Ascona, Switzerland, August 2003.
10. J. M. Kahn, R. H. Katz and K. S. J. Pister, Mobile networking for Smart Dust, *Proc. 5th Annual International Conference on Mobile Computing and Networking* (MOBICOM'99), Seattle, WA, August 1999.
11. M. Reed, P. Syverson, and D. Goldschlag, Anonymous connections and onion routing, *IEEE Journal on Selected Areas in Communications*, 16(4), 1998, 482-494.
12. Perrig, R. Szewczyk, V. Wen, D. Culler and J. D. Tygar, SPINS: Security protocols for sensor networks, *Proc. 7th Annual International Conference on Mobile Computing and Networking* (MOBICOM'01), Rome, Italy, August 2001, 189-199.
13. M. K. Reiter and A. D. Rubin, Crowds: Anonymity for Web transactions, *ACM Transactions on Information and System Security*, 1(1), 1998, 66-92.
14. P. Saffo, Sensors, the next wave of innovation, *Communications of the ACM*, 40(2), 1997, 93-97.
15. A. Scarlata, B. N. Levine, and C. Shields, Responder anonymity and anonymous peer-to-peer file sharing, *Proc. of IEEE International Conference on Network Protocols* (ICNP'2001), November 2001.
16. S. Shenker, S. Ratnasamy, B. Karp, R. Govindan, and D. Estrin. Data-centric storage in sensornets, *Proc. 1st ACM SIGCOMM Workshop on Hot Topics in Networks*, October 2002.
17. C. Shields and B. N. Levine, A protocol for anonymous communication over the Internet, *Proc. of 7th ACM Conference on Computer and Communication Security*, Athens, Greece, November 2000.
18. P. Syverson, D. Goldschlag, and M. Reed, Anonymous connections and onion routing, *Proc IEEE Symposium on Security and Privacy*, Oakland, CA, August 1997, 44-54.
19. P. Syverson and S. Stubblebine, Group principles and the formalization of anonymity, *World Congress on Formal Methods*, Toulouse, France, September 1999, LNCS 1708 Springer-Verlag, 814-833.
20. Wadaa, S. Olariu, L. Wilson, K. Jones, and Q. Xu, On training wireless sensor networks, *Proc. 3rd International Workshop on Wireless, Mobile and Ad Hoc Networks* (WMAN'03), Nice, France, April 2003.
21. Wood and J. A. Stankovic, Denial of service in sensor networks, IEEE Computer, 35(4), 2002, 54-62.