Group Monitoring in Mobile Ad-Hoc Networks

Albana Gaba, Spyros Voulgaris, and Maarten van Steen

Vrije Universiteit Amsterdam
{agaba,spyros,steen}@cs.vu.nl

Abstract. Maintaining bonds of cohesion between members of small groups in densely populated venues (e.g., a family in an amusement park, or some friends in a stadium) is increasingly gaining interest, both as a safety measure against malicious activity and as a convenient tool to prevent group splitting. Note that the use of mobile phones is often ruled out in such scenarios, due to extreme network load. Current solutions are typically based on custom installations of antennas, centralized control, and user devices with high transmission power.

In this work we propose a novel method for anonymously spreading presence information among group members in dense environments, based on a completely decentralized mobile ad hoc network approach. Our system operates independently of any infrastructure and is targetted at resource constrained, inexpensive and expendable user devices. Quite importantly, our system protects the privacy of its users, both for their safety and for ethical reasons.

Keywords: Ad hoc, MANET, group monitoring, presence management.

1 Introduction

Advancements in hardware technologies have led to tiny wireless devices equipped with sensing capabilities, location aware, and able to communicate with each other. These wireless devices have enabled many new applications running in various environments, from military to civilian, that typically monitor events in large scale.

In this paper we focus on the design issues required for building a system that allows people to monitor the presence of each other in crowded areas (e.g., families in an amusement park, friends in a concert). We envision a decentralized system where each person carries a wireless device, collectively forming an ad-hoc network. Group members exchange messages indicating presence information, like location, by relaying them through the ad-hoc network.

The described scenario imposes several requirements. First, sharing confidential data over an untrusted medium, such as the wireless network, may leak information about the persons to third parties. More specifically, adversaries can observe and manipulate the content of packets sent over the network. Even with encryption, though, adversaries may deduce sensitive information by analyzing the traffic of packets and/or by tracing packets from source to destination. In particular, location information of the parties involved in the communication

A.B. Sideridis and Ch. Z. Patrikakis (Eds.): e-Democracy 2009, LNICST 26, pp. 441–451, 2010.

can be deduced. In addition, the routing of packets may be disturbed. For instance, by selectively dropping packets, a targetted group may not receive any fresh messages from a specific member. In case the contents of the message are not well protected, messages may be altered or replayed in a way to mislead group members regarding their peers. For instance, a child that is far from its group, could appear to be in the vicinity.

In addition, the density of the network and the mobility may impact significantly the performance of the delivery of the messages, which may be degraded according to the conditions of the network. The system should be able to work even in harsh environment conditions and tolerate delays and packet loss. The work presented in this paper is independent of the underlying radio technology, as long as it supports message broadcasting. However, low-power radio devices are preferred to more energy-hungry radio technologies, such as Wi-Fi.

In this paper we present our ongoing work on autonomic group monitoring in highly populated areas. We discuss requirements and challenges that derive from building such a system, like issues related to the network characteristics, communication between nodes, etc. In particular, we focus on privacy and anonymity issues concerning group communication.

2 System Model

2.1 Overview

For our application, we assume a crowd of people each carrying a small networked device, such as an electronic badge. Each person belongs to one group and is capable of exchanging messages with its group members. In addition, we assume that there is no pre-installed communication infrastructure. This means that messages should be communicated through the network formed by the crowd as a whole. The main purpose of our application is to allow group members to monitor each other's presence. To this end, each group member periodically transmits presence information such as its relative or absolute location coordinates. This information can then be used by the receiver for further processing. For example, a recipient may conclude that a member is too far away from the others, or may deduce the movement of a group member, and so on. To make this work, we need to meet a number of rather stringent requirements.

These requirements are broadly grouped in the ones related to *privacy* and the *technical* ones, and are laid out in the following two sections.

2.2 Privacy Related Requirements

Privacy plays a central role to our application. In short, group members should share presence information among themselves and with no one else. That is, no data should leak to other nodes, including group member identities as well as the information exchanged.

Anonymity of Nodes. A fundamental aspect of privacy is the anonymity of group members. That is, the identity of the source and recipient(s) of aired messages should not be disclosed to adversary nodes. There are a number of reasons for that.

First, to prevent adversaries from tracking down individual persons. If an adversary is capable of recognizing the identity of the person who issued a message, he can use our application to track down that person based on the messages he is sending. This is crucial for safety. For instance, it may be possible to detect when group members are isolated from the rest of the group, and therefore more vulnerable. This is vertically opposite to the goals of our application, which aims at protecting persons from malicious activity.

Second, to prevent attacks directed at specific groups. By being able to read the sender or receiver identifiers in aired messages, an adversary could *selectively* tamper with the messages of a specific group. For instance, he could drop, delay, or corrupt messages of a particular group, or he could replay them with wrong information.

Finally, node anonymity is important for keeping group composition undisclosed to third parties. Exposing the sender and receiver identities in traveling messages could unmask group membership, by revealing the sender/receiver relationships.

It should be clear from the above that messages should not lead back to any node (sender/receiver) or group identifier. Therefore, source and destination nodes should remain anonymous.

Message Authenticity. It should not be possible for an adversary to *impersonate* a legitimate person, that is, generate messages that appear to be coming from that person. Consider a scenario where a member of a group is drifting away, while an adversary's instrumented device (falsely) assures the other members of the group that their friend is nearby. This would annul the principal goal of our application.

Thus, recipients of a message should be able to confirm the authenticity of the message. This implies some sort of encryption regarding the sender identity, as we will see in Section 4, that prevents adversaries from impersonating legitimate users. As it turns out, the goals of node anonymity and authenticity of messages are closely related.

Message Content Encryption. The information conveyed in the exchanged messages should not be exposed to adversaries. Similarly to message authenticity, this requires some sort of encryption of the messages, to avoid *eavesdropping*. If the content of the messages is not protected, then eavesdroppers can monitor the group members as if they were group members.

2.3 Technical Requirements

Here we list the technical requirements for our system.

Independence from Any Infrastructure. One of the basic requirements for our system is to operate in an autonomous, self-contained manner, not dependent on any type of custom infrastructure or telecommunication providers.

The system should be usable anywhere, without requiring the preinstallation of any type of infrastructure. Consider, for instance, a school trip to the countryside, a company picnic, a family visiting a crowded beach, or other outdoor activities. Financing the installation of infrastructure at such places would be a serious issue for the deployment of our system.

It can be argued, however, that in most developed countries GSM and GPRS coverage spans the whole land, including distant areas in the countryside. Even if we take GPRS for granted, the cost involved with it, as well as the need to have one mobile phone per monitored entity (e.g., children, pets, etc.) may be a restrictive factor. This becomes more clear in the scenario of a trip to a foreign country. Current roaming charges would form a prohibitive barrier to the frequent (once every few seconds) reporting of presence information. Additionally, our system should also work in very crowded areas, like a concert, where the capacity of the existing GSM infrastructure is usually exceeded.

Node Mobility and Density. The application is required to operate in highly mobile environments. Mobility is inherent in the scenarios we are targetting, such as amusements parks, shopping centers, etc.

Additionally, our system should operate well under high concentration of people, where infrastructure based networks typically run out of capacity.

Unobtrusive Devices. We target at tiny, unobtrusive devices (e.g., bracelets) that can be easily carried by anyone, including children or even pets. In addition, the devices are required to be inexpensive as this would allow them to be massively affordable. The small size of the devices imposes a serious limitation on their battery capacity, so minimal energy consumption is crucial to their sustainability. As a consequence, handheld devices such as smart phones, PDAs, etc., that rely on Wi-Fi medium access, are ruled out. The high energy consumption of Wi-Fi interfaces presents a significant barrier in building tiny, unobtrusive devices that keep communicating for a suffiently long duration (e.g., a full day or more).

A representative architecture that fits the aforementioned requirements is the resource-constrained, low-power architecture of the Berkeley nodes [1]. Equipped with an 8 MHz micro-controller processor, 10kB of RAM, 1 MB of external flash memory and a 256 kbps data rate radio, these devices can compute simple operations in an energy-efficient manner. They have shorter communication range, compared to Wi-Fi devices, that can reach up to 100 meters, but this is not an obstacle for the connectivity in our network as we assume highly populated networks.

Timely Delivery. It is important that people receive timely updates about the state of their group members. This implies that the ratio of undelivered packets should be kept low, and should gracefully degrade with the distance of group members. Furthermore, if the network reliability degrades, it should be detectable, allowing people to rely on other means for monitoring each other.

3 System Design

In this section we discuss the impact of the application requirements on the design decisions.

We consider the following *adversary model*. One or more adversaries eavesdrop the messages delivered through the wireless medium and, when necessary, move from one place to another. The adversaries, equipped with powerful devices, are capable of computing with high probability the direct (one-hop) sender of a message. To this end, special electronics, such as directional antennas and spectrum analyzers, can be used to compute the angle of arrival and the received signal strength of a message and infer its direct sender. Successively, in order to efficiently locate and track people, the adversaries may share this information with each other in a private way using long-range wireless communications and correlate the overheard data.

3.1 No Node Identifiers

Communication between nodes should not undermine the privacy of the people involved. Based on the messages nodes exchange, third parties should not be able to detect locations, motion patterns, or any data related to the nodes. Moreover, sender-receiver relationships, that is, the composition of groups, should not be revealed. This would allow adversaries to trace target people or even groups from distance.

Along these lines and in order to safeguard user anonymity, one of our design choices is to prevent node identities from appearing unencrypted in messages. Doing so would constitute a major threat to user anonymity, as explained below.

If the identity of nodes can be revealed to third parties, then their messages may be traced and therefore the privacy of individual nodes may be violated. Typically, a source node includes its ID in the clear in its messages so that they are recognized by the receivers. However, an adversary node can exploit this feature, and by standing nearby a target node it can figure out the ID of that node. Thereafter, this adversary, or a network of collaborating adversaries, could track that node in space and time by following the flow of its messages in the network. Additionally, they could selectively *suppress*, *corrupt*, *delay*, or *replay* at a different time messages of given nodes. Even worse, if the sender ID is accessible and the data is not protected well enough, an adversary could *impersonate* a legitimate user, sending misleading messages to its group partners. For instance, if location information is included in messages, an adversary could make a node appear to be nearby while it is far away, by impersonating it, or by replaying its old messages.

Group composition, that is, the set of source-destination relationships, should remain undisclosed to third parties. If it is possible to identify the source nodes of the messages (as seen before) and the nodes that compose a group (*communication patterns*), even if the exchanged data is not accessible, multiple adversaries can analyze the traffic of messages on the paths that link source nodes to destinations and monitor the locations and the movements of the whole group. As a

consequence, it may be possible to detect when group members are isolated from the rest of the group. Moreover, similarly to what said about single nodes, the exchanged messages may be captured and suppressed selectively, but at a group level. For example a determined node can be isolated in a way that its messages are dropped for all the destinations. On the other hand, if the content of the messages is not protected, then eavesdroppers can monitor the group members just as internal group nodes.

3.2 No Point-to-Point Routing

The restriction on including unencrypted node IDs in messages has a direct impact on the applicability of classic routing protocols. Most point-to-point routing protocols operate based on addresses, and, at a minimum, the *recipient address* of a message should be readable by any intermediate node. As such, address-based point-to-point routing protocols are deemed unsuitable for our application.

Some point-to-point routing protocols, such as geographic routing [7], operate without using explicit node addresses. Even these protocols, however, are inappropriate for our application. For instance, in the case of geographic routing, disclosing the location of the destination—even without its address—is sufficient to reveal the actual person; let alone the difficulty of knowing a person's exact location to send it a message.

In the generic case of point-to-point routing, a team of collaborating adversaries could track a message from its source to its destination, inferring the group membership. Figuring out that a node is the source of a message is trivial to accomplish by capturing all messages broadcast around the node's location. If the message is first broadcast by the node, then this node is the source of the message. If it has been heard before, then this node is just forwarding that message. Similarly, for inferring the recipient, the adversaries could follow the propagation of a message until the last time it is forwarded. Unless the message was lost or some other routing problem occurred, the recipient is within one hop of the last node forwarding the message.

In addition to the privacy related issues preventing the use of routing, node mobility and network size further strengthen the arguments against it. Routing in dynamic networks is not very efficient, and often involves a large overhead in the type of floods to (re-)discover routes to relocated nodes.

3.3 Gossip-Based Message Propagation

A key design choice in our application is the use of gossiping to propagate messages. This is favored by a number of reasons in addition to the unsuitability of routing protocols discussed in the previous section.

First, gossiping is ideal for dense networks, as it completely eliminates the need for flooding. Each node broadcasts messages at a steady period, avoiding transmission bursts. Second, redundancy is an inherent property of gossiping. Redundancy is crucial for multihop communication in dynamic networks, such as the ones we are targeting here. This is particularly important in highly mobile

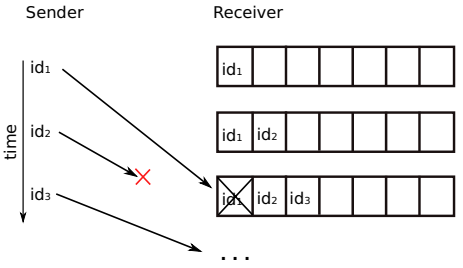

Fig. 1. Window of expected IDs

environments. Third, gossiping propagates a message without revealing its recipient. Finally, gossiping constitutes a very simple technique for disseminating information in a fully decentralized fashion, which makes it appealing for our large network of resource-constrained devices.

4 Anonymous Communications

Our goal is to design a system that allows nodes to send messages in a way that only the intended recipients are able to observe and decipher. Therefore, we propose the following protocol.

During its lifetime, each node uses IDs out of a custom *secret sequence* to mark its messages. The ID space is long enough to guarantee a unique ID per message, among contemporary messages, with very high probability. To establish an one-way communication between two nodes, the sender S shares in advance its secret sequence of IDs with the receiver R, and the two nodes synchronize their clocks. Given that messages are transmitted periodically, with a known period, the receiver R is in a position to know at any given moment what message IDs to expect in S's messages, as it can estimate which IDs from the sequence S used to mark its last few messages.

Under ideal conditions (i.e., no message losses, no delays), the receiver R would get the expected IDs (messages) from S on time and in the order they were sent. In practice, however, messages do get dropped, or experience variable delays otherwise, due to multi-hop propagation. To cope with network unreliability, R keeps a window of *expected* IDs from S. As shown in Fig. 1, at every interval R inserts the next ID from S's sequence to the window, which is also the ID S uses to mark its new message. As messages arrive, the IDs in the respective window are looked-up, and if matched, the message is known (with very high probability) to have originated at S. An ID (message) is considered to be lost when it does not arrive within a maximum time T. The window size may be dynamic depending on the latency of the network.

Rather than storing whole sequences of IDs, nodes use iterative functions that generate such sequences on the fly. In particular, we adopt Pseudo Ran-

dom Number Generators (PRNG) for that goal. A PRNG is an algorithm for generating a sequence of numbers that approximates the properties of random numbers. The sequence is completely deterministic and can be reproduced by using the same set of initial values, called the PRNG's seed. The characteristics of PRNGs turn out to fit our requirements for sequences of one-time IDs, as shown below:

- *unpredictable*, by observing a sub-sequence of one-time IDs originated by a specific node it is hard to predict the successive ones,
- *unobservable*, by observing any sub-sequence of one-time IDs it is not possible to draw any conclusion about relationships between IDs,
- *reproducible*, it is possible to reproduce the same sequence of IDs by other nodes,
- *simple*, given the constrained resources of the nodes, it is required that inexpensive algorithms are employed to generate the sequence of IDs.

The result of employing a one-time ID for each message is that no information, like the sender or recipient identity, can be deduced by eavesdroppers who analyze the traffic of messages or trace them hop-by-hop. Exception is made when the adversary is within the communication range of a sender and can directly observe its messages. But even in this case, if adversaries trace the identified message, it is not possible to find the recipient. Also, since the successive messages have different IDs there is no new information gained by capturing one-hop messages.

4.1 Discussion

Pseudo Random Number Generators exist in a wide range and differ in complexity according to the purpose they are used for. For lack of space we will see only two of them. The simplest ones are the Linear Congruential Generators, mainly used for simulations. Successive values are computed by $id_{k+1} = (A{*}id_k{+}B)\ mod\ N$, where A, B, and N are fixed parameters, and id_0 is the seed. LCGs have shown to be predictable when *enough* consecutive output values are observed [3]. A more complex PRNG is Blum Blum Shub [2]. Successive values are computed by $id_{k+1} = id_k^2\ mod\ N$, where N is a fixed product of two large primes and id_0 is the seed. BBS provides strong guarantees of unpredictability, but is slower than LCGs, as it requires that $log_2 n$ bits are extracted from each id_k. According to the characteristics of the system one may choose the PRNG that is best suitable.

Authentication of Messages. A node R accepts every incoming message whose ID matches the window of expected IDs of node S. In this context, we should consider the case that R might validate messages generated by other nodes, but whose ID happens to coincide with some entry in the window. To this end, the ratio between the ID space and the network size should be set appropriately in order to minimize the risk of collisions between IDs. But even when an ID collision occurs, it only affects an isolated message rather than the whole series of messages from a given sender.

Data Protection. As a result of employing one-time IDs, we obtain messages unobservable to anyone but the group members. This means that no correlation can be made between the overheard messages. Therefore, according to the level of confidentiality we want to achieve, if the exchanged information does not lead to any specific node in the network, we may opt to use inexpensive encryption algorithms (e.g., stream cyphers), if at all.

Alternative Solution. Using encryption would be an alternative to our protocol. Before sending a message, nodes would encrypt it entirely with a symmetric key shared with the intended recipients, making it unreadable for anyone else. However, a node would have to attempt to decrypt *every* message it sees, in search of decipherable messages encrypted by its friends. Although a viable solution, it appears to be more complex compared to our PRNG based approach. As we are striving for low power and resource constrained devices, we opt to employ the PNRG approach. Nevertheless, it is left as future work to assess the efficiency of each approach, particularly when the algorithms are implemented in hardware.

5 Related Work

The problem of node location privacy as a consequence of traffic analysis and packet tracing has increasingly gained the attention of the research community. In [4], Deng et al address the problem of hiding the location of the receiver (i.e., base station) in a sensor network. Assuming that the traffic of the whole network is directed towards the base station, they concentrate on protecting the system from traffic analysis through measurements of traffic rates at various locations. Randomized routing and fake message injection are introduced to prevent an adversary from locating the network sink based on the observed traffic patterns. The same problem of hiding the sink location in sensor networks is addressed in [5], with the difference that they focus on packet tracing attacks, which come as result of multi-hop deliveries from sources to the sink. In [6], the problem of source location privacy is studied in a sensor network. The adversary traces back the packets received at the base station, hop-by-hop up to the source. The authors introduce *phantom routing* that consists of a combination of random walk with flooding/shortest-path routing when the message approaches the base station.

In [9], the impact of anonymity as a result of mobility is studied and some anonymity-preserving routing protocols are compared. Anonymity in MANETs has been studied in ANODR [8], MASK [12], RIOMO [10], ARM [11]. These works concentrate on routing on demand in mobile ad-hoc networks. The basic idea behind anonymous routing is that nodes keep an entry for each anonymous RREQ they forward in order to recognize it on the way back (RREP) and forward it to the right neighbor. The network and computation assumptions these protocols make are prohibitely expensive to be adopted in resource-constrained networks. For instance, in ANODR, a fresh public/private key pair is generated for each RREQ by the forwarding nodes, and messages grow in size as they are

forwarded along the path. Along these lines, the other protocols assume encryption/decryption of every RREQ and RREP message forwarded. In [12] large storage is required to keep a set of pseudonyms. In addition, the aforementioned protocols make assumptions such as limited mobility, reliable communication, and symmetric links, which are not realistic in our model.

6 Conclusions and Future Work

In this paper we sketched the design of a system for monitoring group members in a mobile ad-hoc network. After a first consideration of the factors that charachterize the system, we pointed out the possible constraints and challenges that such system might face. Particular focus is devoted to the privacy of nodes involved in these communications, such as location and sender-receiver identity. We proposed an anonymity scheme where nodes use anonymous messages that do not lead to any information concerning the communicating parties or related to other messages. The ultimate goal of the proposed system is to ensure unobservability of the messages exchanged between group members while adding low overhead in computation resources, storage or message size. In the future we propose to implement such system and run experiments on nodes to assess the efficiency of our protocol.

References

[1] `http://webs.cs.berkeley.edu/tos/`
[2] Blum, L., Blum, M., Shub, M.: A simple unpredictable pseudo random number generator. SIAM J. Comput. 15(2), 364–383 (1986)
[3] Boyar, J.: Inferring sequences produced by pseudo-random number generators. J. ACM 36(1), 129–141 (1989)
[4] Deng, J., Han, R., Mishra, S.: Countermeasures against traffic analysis attacks in wireless sensor networks. In: SECURECOMM 2005: Proceedings of the First International Conference on Security and Privacy for Emerging Areas in Communications Networks, Washington, DC, USA, pp. 113–126. IEEE Computer Society, Los Alamitos (2005)
[5] Jian, Y., Chen, S., Zhang, Z., Zhang, L.: Protecting receiver-location privacy in wireless sensor networks. In: INFOCOM, pp. 1955–1963 (2007)
[6] Kamat, P., Zhang, Y., Trappe, W., Ozturk, C.: Enhancing source-location privacy in sensor network routing. In: Proceedings of the 25th IEEE International Conference on Distributed Computing Systems, June 2005, pp. 599–608 (2005)
[7] Karp, B., Kung, H.T.: Gpsr: greedy perimeter stateless routing for wireless networks. In: MobiCom 2000: Proceedings of the 6th annual international conference on Mobile computing and networking, pp. 243–254. ACM, New York (2000)
[8] Kong, J., Hong, X.: Anodr: anonymous on demand routing with untraceable routes for mobile ad-hoc networks. In: MobiHoc 2003: Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing, pp. 291–302. ACM, New York (2003)
[9] Kong, J., Hong, X., Sanadidi, M.Y., Gerla, M.: Mobility changes anonymity: Mobile ad hoc networks need efficient anonymous routing. In: IEEE Symposium on Computers and Communications, pp. 57–62 (2005)

[10] Rahman, S.M.M., Nasser, N., Inomata, A., Okamoto, T., Mambo, M., Okamoto, E.: Anonymous authentication and secure communication protocol for wireless mobile ad hoc networks. John Wiley & Sons, Ltd., Chichester (2008)
[11] Seys, S., Preneel, B.: Arm: Anonymous routing protocol for mobile ad hoc networks. In: AINA 2006: Proceedings of the 20th International Conference on Advanced Information Networking and Applications, Washington, DC, USA, pp. 133–137. IEEE Computer Society, Los Alamitos (2006)
[12] Zhang, Y., Liu, W., Wenjing, L.: Anonymous communications in mobile ad hoc networks. In: INFOCOM 2005: 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE (2005)