Privacy Challenges in Delay-Tolerant and Restricted-Route Networks

Joss Wright
Oxford Internet Institute
1 St. Giles
Oxford. UK
OX1 3JS
joss.wright@oii.ox.ac.uk

Ian Brown
Oxford Internet Institute
1 St. Giles
Oxford. UK
OX1 3JS
ian.brown@oii.ox.ac.uk

ABSTRACT

We examine privacy-preserving protocols in the context of delay-tolerant networks, with particular application in sparsely connected environments where traffic is routed via predictable routes. We consider potential threats and attacker models against the expected usage of such networks, and explore a range of technologies that aim to support privacy-preservation in applications of significant interest to users in these environments.

Categories and Subject Descriptors

C.2.0 [**Computer-Communication Networks**]: General—*security and protection*; C.2.4 [**Computer-Communication Networks**]: Distributed Systems—*distributed applications*; H.3.4 [**Information Systems**]: Systems and Software—*distributed systems*

Keywords

Anonymity, Privacy, Security, Distributed Systems, Delay-Tolerant Networking

1. INTRODUCTION

Delay-tolerant networking (DTN) [2] is a development from work designed for interplanetary networking, in which extreme communication delays and transmission disruptions are significant factors. Delay- and disruption-tolerant networking has increasingly seen use in more terrestrial environments that experience similar networking constraints, such as vehicular ad-hoc networks and rural or other sparsely-connected environments [8].

The delay-tolerant approach allows networked services to function where communications are often dropped or highly delayed, which limits the interactivity of protocols and, to some extent, traffic volumes. The constraints under which such networks function have severe effects on many security and privacy-enhancing protocols.

ExtremeCom 2010 Dharamsala, India

In networks for which achieving connectivity is itself a major challenge, it could be argued that concerns of user security and privacy are of low importance, and further that environments with poor technological infrastructures are less likely to contain malicious attackers. We reject this view. Connectivity in rural environments can communicate medical data, report critical news, and enable e-voting and other forms of representation and participation; any one of these provides motivation for consideration of security and privacy, even in severely resource-constrained networks.

2. GOALS

We aim to support privacy within the constraints of delay-tolerant networking, through techniques for traffic anonymisation and data-level privacy. In contrast to overlay protocols designed for the general Internet, we seek to exploit the most effective privacy-preserving techniques possible based on the nature of data, and identify those situations in which privacy is likely to be unobtainable.

In general, applications that can support higher latencies are more amenable to traffic obfuscation; streams of information, such as voice-over-IP, that require close to real-time communication are less appropriate for anonymisation, and any further delay beyond that imposed by the DTN architecture itself is undesirable.

We expect that any protocols for enhancing user privacy should be robust and transparent to end users, and should not rely on users to make regular decisions regarding privacy and security. Selection of privacy properties should therefore be largely automatic at the protocol level.

Towards these goals, we take advantage of potentially privacy enhancing properties of mobile delay tolerant platforms, such as the broadcast nature of radio channels; the ability to exchange data directly with nearby peers; and the potential for nodes and peers to move between networks.

2.1 Expected Environment

The network infrastructure that we consider consists of a number of relay base stations that direct traffic between themselves, supporting a localised wireless network offering connectivity to clients. Traffic is routed from these relays to a small number of Internet gateways as shown in Figure 1.

Each base station provides access to the wider Internet for

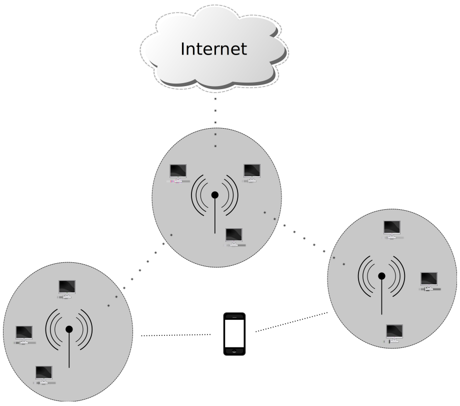

Figure 1: Separated wireless zones connected by antenna-to-antenna links; mobile devices roam between zones.

local clients, which also communicate amongst themselves via a local mesh network. Truly mobile devices, such as mobile phones, should be able to travel between these localised zones and make use of the base-station relay when in range.

The elements of the network are therefore:

- **Uplink:** A connection to the wider Internet that may itself be intermittent or slow. From the point of view of a downstream client this is a severe potential threat to availability, security and privacy if untrusted.
- **Relay Base Station:** Provides connections to other base stations and, ultimately, the wider Internet. Local clients connect wirelessly to the base station for all routing beyond locally-reachable peers.
- **Local Client:** Standard user machines that take part in a local mesh network around a given base station. These can be single-user machines or can take the form of a public station. Some local clients may be in a position to connect to more than one base station.
- **Mobile Device:** Low-powered devices, in terms of processing speed and battery life for communications. These devices, more than any other, may roam between various base stations.

3. CHALLENGES

There are several assumptions typically made by privacy-enhancing technologies that are, to a greater or lesser extent, broken by the environment described above. One of the most significant of these is the lack of network-wide consistency for shared data, which severely limits the effectiveness of traditional public-key infrastructures as clients are unable reliably to verify certificates, particularly through online checks for freshness of certificates, or to obtain public keys on-demand from a trusted authority.

The lack of a global view of the network has other disadvantages for traditional anonymity protocols. To gain an even distribution of messages across the network, and thus to provide the maximal level of confusion as to the source and destination of a given message, anonymity protocols typically require any participant in the network to communicate directly with any other participant. This seeks to defeat network partitioning, and to require attackers to maintain a view of a large portion of the network in order to launch attacks [5].

Any protocols used across high-latency connections must be as non-interactive as possible, as each additional reply greatly increases the time taken to complete a protocol. Traditional cryptography will be feasible between local clients, allowing for local broadcast and message passing, however messages travelling longer distances across the network will be largely constrained to long-term pre-shared keys and similar approaches.

Another technology of use is identity-based encryption (IBE) [1], that makes use of a plaintext identifier as the public key used to encrypt messages. The corresponding decryption key can be obtained by the recipient from a trusted key server, and thus reduces the need for a network-wide key infrastructure. Identity-based encryption provides partial solutions to some of the basic cryptographic constraints of a delay-tolerant network [6], but is both resource-intensive and largely patent-encumbered. Further, the inability of a participant to know the overall state of the network raises its own problems with identity-based encryption, in that collisions between namespaces for public key identifiers become possible, potentially leading to conflicting public and private keys.

Attackers in the network architecture above are themselves constrained in their inability to view the entire network, and thus do not fit into the standard view of the *Dolev-Yao* attacker that has all the capabilities of the network itself. The major threat, in the environments that we consider, comes in the form of relay operators on the fixed route of a given client. All traffic from a client to the wider network must pass through these relays, and as such their view of users local to their antenna is complete. As traffic is unlikely to be routed through remote network end-points by other clients, anonymisation of messages is largely restricted to those clients that share a single relay or path through the network. This problem is increasingly severe the more remotely located the given antenna.

4. SOLUTIONS

The problems identified above for the preservation of privacy and security are severe. We briefly present here a number of approaches that have the potential to contribute towards protecting user privacy in networks of this form.

4.1 Opportunistic mixing

The most popular approach to anonymising traffic is the *mix*, originally presented by Chaum [3] to provide sender and recipient unlinkability for email. Briefly stated, a mix is a store-and-forward node that accepts a number of encrypted messages until a given threshold criterion is reached, at which point it forwards stored messages in random order. Messages are cryptographically transformed by the mix in order to prevent an observer from trivially linking ingoing and outgoing messages. To reduce trust in any single server, mix-based approaches typically route messages through a random selection of intermediary mixes before their destination. The mix, under certain strong assumptions, provides computationally secure resistance to traffic analysis.

The mix approach imposes a delay on traffic while the pool of messages is collected, and as such is largely inappropriate for realtime traffic. The approach of delayed store-and-forward message passing for the purposes of anonymity has the potential, however, to interact positively with the underlying store-and-forward nature of the delay-tolerant networking architecture.

We suggest that each mesh network surrounding a relay run a localised mix network, and that messages passing through that relay should be pushed to the local network for mixing before being re-injected into the wider network through the relay. This approach allows messages from outlying relays to be mixed with messages from the entire path along which they travel.

There are severe limitations to this approach when compared to an ideal mix network, and the relative level of trust required by each relay is greatly increased as clients may have little or no choice in the path taken by their messages. One major issue is to ensure that malicious relays pass messages out to local networks for mixing rather than directly forwarding, or simply dropping, them.

4.2 Prioritisation

The DTN architecture provides a prioritisation scheme that allows a participant in the network to inform routers of the relative urgency of its own packages. This prioritisation applies on a per *endpoint identifier* (EID) basis, with multiple clients able to share an EID for multicast or anycast communications [7]. Clients are unable to affect the relative prioritisation of their traffic against other clients in the network, but are able to indicate that certain of their message bundles are of a lower delivery-priority than others. We examine the potential of multicast identifiers later, but observe that the coarse-grained per-client prioritisation of bundles could be employed to signify message bundles that are appropriate for privacy technologies with higher or lower latency requirements.

4.3 Broadcast

The assumption that the majority of traffic in the proposed environment will be transmitted wirelessly suggests that the broadcast nature of the medium could be exploited to obfuscate the origin of messages amongst all senders. In practice, exploiting broadcast would largely require the hiding of client identifiers through techniques such as MAC spoofing and dynamically assigned IP addresses.

Of particular application in delay-tolerant networking is the concept of multicast endpoints described by the DTN specification. As mentioned above, an *endpoint identifier* (EID) is a uniform resource identifier used to identify some subset of the network for the purposes of routing or ownership of message bundles. The DTN specification allows for *multicast* and *anycast* identifiers in which the EID refers to all members of the set or, respectively, a single member from that set. The apparent intention of this is to allow for robust content delivery, allowing for message broadcast or for messages to be considered successfully delivered even if only a single recipient from a group is available, but it has the potential to provide a mechanism through which to obfuscate the actions of individuals.

4.4 Content Distribution

The mirroring and sharing of content has the potential to be both bandwidth-efficient and privacy-preserving. For certain classes of information, such as news sources, we anticipate that content could be replicated across a locality. This approach would improve availability for clients with lower connectivity, reduce direct load on the relay through peer-to-peer sharing and obscure observable relationships between clients and the news sources that they individually consume.

Content distribution of this form also lends itself to anonymous message-board systems, a form of naive private information retrieval [4], in which a set of appropriately encrypted messages are broadcast from the relay to each client in the local mesh. Given the set of all messages, a client with the appropriate key can decrypt their own messages without revealing to the relay which messages it has accessed.

4.5 Roaming Devices

Mobile phones that can roam between localities provide an unreliable, and potentially very slow, channel beyond the normal route for message bundles. This channel is unlikely to be of use in standard applications, however the lack of reliance on the local antennas makes roaming devices a potential side channel that could be used for low-bandwidth and one-off purposes, such as publication of cryptographic key information or reporting of local relay behaviour.

It should also be noted that the diagram shown in Figure 1 does not admit the possibility that a local client is within wireless range of multiple antennas. That client could function as an occasional bridge between two localities, functioning as a more reliable equivalent to a roaming mobile device.

5. CONCLUSIONS

There are many issues in providing secure and privacy preserving applications to users of delay-tolerant and route-restricted networks. It is clear that traditional security approaches are not, in general, applicable to these environments. Of particular note is the additional requirement of a minimization of communication, both in terms of the amount of traffic generated and the number of steps, in protocol design.

The approaches laid out in this paper are a first step towards adapting traditional privacy and anonymity protocols to function in, and ideally take advantage of, the environment we have described. The specifics of how to make these protocols work under the constraints of delay-tolerant and restricted-route networking, and the attack models that arise, raises a number of interesting questions for future research.

References

[1] D. Boneh and M. Franklin. Identity-based encryption from the weil pairing. *SIAM J. Comput.*, 32(3):586–615, 2003.

[2] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss. Delay-Tolerant Networking Architecture. RFC 4838 (Informational), Apr. 2007.

[3] D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2), February 1981.

[4] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan. Private information retrieval. *Foundations of Computer Science, Annual IEEE Symposium on*, 0:41, 1995.

[5] R. Dingledine, V. Shmatikov, and P. Syverson. Synchronous batching: From cascades to free routes. In *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)*, volume 3424 of *LNCS*, pages 186–206, May 2004.

[6] A. Kate, G. M. Zaverucha, and U. Hengartner. Anonymity and security in delay tolerant networks. In *Security and Privacy in Communications Networks and the Workshops, 2007. SecureComm 2007. Third International Conference on*, pages 504–513, 2007.

[7] K. Scott and S. Burleigh. Bundle Protocol Specification. RFC 5050 (Experimental), Nov. 2007.

[8] T. Spyropoulos, T. Turletti, and K. Obraczka. Routing in delay-tolerant networks comprising heterogeneous node populations. *IEEE Transactions on Mobile Computing*, 8(8):1132–1147, August 2009.