Mobile Hybrid Anonymity Network

Francesco Pittaluga
University of Florida
f.pittaluga@ufl.edu

Abstract

Recent events, such as the technologically driven persecution of political activists and the revelations of mass spying programs across the western world, have heightened the need for mobile anonymous communication networks that do not rely on a trusted network operator. Such anonymous networks can be achieved by combing Wi-Fi and Bluetooth-based peer-to-peer (P2P) networks with cellular networks. Wi-Fi and Bluetooth-based P2P networks operate independent from cellular networks. Thus, in theory, nodes in these P2P networks can achieve crowd like anonymity when sending packets through a cellular network. In this paper, we combine P2P networks, cellular networks, to develop a hybrid Crowd-mobile anonymity network that preserves anonymity despite a compromised cellular network.

1. Introduction

In 1998, Reiter and Rubin introduced the Crowds for anonymous web transactions [3]. Nodes in a Crowd achieve anonymity by forming virtual "message forwarding" circuits to hide the true source of a web request. Recently, [1] proposed a crowd-like anonymity network that leverages mobile phones P2P and cellular capabilities. Like crowds, nodes (mobile phones), in this network, achieve anonymity by forwarding messages among themselves to hide the true source of a web request. The key contribution is that messages are forwarded via P2P networks setup by the nodes, so the the cellular network operator cannot by default observe these exchanges.

In this paper we build on the approach proposed by [1] and in the process address the following weakness of the approach: 1) Nodes are susceptible to position attacks when involved in transactions with multiple requests-responses; 2) The protocol requires source nodes to forward each of their messages to k different peers, which makes them susceptible to predecessor attacks; 3) On the web server side, the proposed protocol does not conform to current web standards. Our proposed protocol is not susceptible to position attacks, reduces the effectiveness of predecessor attacks, and conforms to standard web server protocols.

More specifically, the contributions of this paper are:

1. Novel hybrid peer to peer and cellular anonymity protocol.
2. Analysis of load cost of the proposed protocol.
3. Analysis of anonymity afforded by the protocol for various attacker models.
4. Simulation of the proposed protocol.

2. Protocol

Let $\mathcal{P}$ denote the set of peers in a p2p network; let $u \in \mathcal{P}$ denote the initiator of message m with target web server s; and let d_i a set of dummy messages. Furthermore, let the number of offspring generated by the K peers in the message tree $p_{i=1...K-1}$ be a iid random variables denoted by ξ_i; let the number of dummy follow-up request a peers $p_{i=1...K-1}$ send to server s be a iid random variables denoted by δ_i; and let the number of offspring of the initiating peer be a random variable denoted by N. Then, the proposed protocol is the following:

REQUEST $(u \rightarrow s)$

User $u \in \mathcal{P}$

1. Generate N
2. With random delay, send message m to s through a cellular network.
3. **for** $i = 0$ to N
   - Generate dummy request d_i
   - Randomly select a peer p_i from $\mathcal{P}$
   - With random delay, send d_i to p_i through a P2P network

Peer $p_i \in \mathcal{P}$

Upon receiving packet d_i from a peer.

1. Generate ξ_i and δ_i
2. With random delay, send d_i to s through a cellular network.
3. **for** $j = 0$ to ξ_i
   - Generate new dummy request d_j
   - Randomly select a peer p_j from $\mathcal{P}$

- With random delay, send d_j to p_j through a P2P network

RESPONSE $(s \rightarrow u)$

User $u \in \mathcal{P}$
Upon receiving response for request m from server s.

1. **if** further requests desired
   - Send next request to server s through cellular network

Peer $p_i \in \mathcal{P}$
Upon receiving response a response from server s for this session V, i.e. the session initiated by request d_i

1. **if** # of dummy request sent for session $V < \delta_i$
   - Generate new dummy request
   - Send new dummy request to server s

---

3. Load Cost

The load costs for the cellular network is equal to the total number of peers in a message tree since each peer in the tree sends a single message to the server through the cellular network. Let K be a random variable that denotes the total number of peers in a message tree. Then, the amortized load cost per message is given by the expected value of K. Message trees are initiated when a single peer sends an initiating message to $N > 1$ peers. Thus, by the branching property of Galton–Watson Processes [2], we can model the growth of the message tree as a Galton–Watson Process X_n, where X_n denotes the total number of peers in generation n and $X_0 = N$.

More formally, let the number of offspring generated by peers $i = 1...X_{n-1}$ in generations $n = 1...\infty$ be a iid random variables denoted by ξ_i^n. Then, the process $\{X_n\}$ evolves according to the recurrence formula

$$X_{n+1} = \sum_{i=0}^{X_n} \xi_i^n. \tag{1}$$

and the expected value of X_{n+1} is given by

$$\mathbf{E}(X_{n+1}|X_0 = N) = N\mathbf{E}(\xi)^n \tag{2}$$

It follows then that the total number of peers in the message tree is given by

$$K = 1 + \sum_{n=0}^{\infty} X_n \tag{3}$$

and that the expected value of K

$$\begin{aligned} \mathbf{E}(K|X_0 = N) &= 1 + \sum_{n=0}^{\infty} \mathbf{E}(X_n) \\ &= 1 + \sum_{n=0}^{\infty} N\mathbf{E}(\xi)^n \\ &= 1 + \frac{N}{1 - \mathbf{E}(\xi)} \end{aligned} \tag{4}$$

converges if and only if $\mathbf{E}(\xi) < 1$. Incorporating the fact that N is also a random variable, the amortized load cost is given by

$$\begin{aligned} L(f_\xi, f_N) &= \mathbf{E}(\mathbf{E}(K|X_0 = N)) \\ &= 1 + \frac{\mathbf{E}(N)}{1 - \mathbf{E}(\xi)} \end{aligned} \tag{5}$$

where f_ξ and f_N denote the probability mass functions of N and ξ respectively.

From equations (4) and (5), we see that the condition $\mathbf{E}(\xi) < 1$ is necessary to keep the load cost bounded. Furthermore, the two degrees of freedom f_ξ and f_N enable trading off tree breadth for depth and vice-versa while preserving the same amortized load cost. The privacy implications of this trade-off are examined in the following section.

4. Anonymity Set

In this paper, we consider three adversaries: the network operator, colluding peers, and the two working together. For all cases we assume an initiating message m, intended for target server s, was sent at time t, and a fixed amortized load cost C. By fixing the load cost, we can examine the trade-offs between the two degrees of freedom afforded by the protocol f_ξ and f_N, respectively the probability mass functions of N, the size of the first generation, and ξ_i^n, the number of offspring generated by peers $i = 1...X_{n-1}$ in generations $n = 1...\infty$. This trade-off is highlighted by equation (5).

4.1. Network Operator

The network operator is a global passive observer of the cellular network channels. Thus, the operator can see all requests made to server s, which enables it to, with probability 1, correctly reduce the size of anonymity set A to

$$A = K + M \geq N + 1 + M \tag{6}$$

where K denotes the number of peers in the message tree, N denotes the number of peers in the first generation of the message tree, and M denotes the number of peers not in the message tree, that sent a request to server s at times $t_{1...M}$, such that $|t_i - t| < \epsilon$.

Since we cannot control the M, let us disregard its contribution to A. Then, the anonymity set size, A, is bounded solely by N, the number of peers in the first generation of the message tree. Thus, examining equation (5) and recalling that we have fixed the load cost to C, we see that the expected value of the anonymity set is maximized when f_N is set such that $E(N) = C$ and $f_\xi = 0$.

4.2. Colluding Peers

Unlike the network operator, colluding peers cannot reduce the anonymity set with any certainty. They can however gain some probabilistic insight into the identity of the initiator by launching a predecessor attacks. The key to the predecessor attack is that the condition $E(\xi) < 1$ must be met in order for the load cost of the network to remain bounded. Assuming this condition to be true, then if one or more colluding peers receive initiating messages, for the same message, from the same peer p_0, then, compared to other peers, their is a higher probability that that p_0 is the initiator. Furthermore, the probability that p_0 is the initiator increases as the number of messages they receive from that peer increases.

Since we're only concerned with messages originating from the message initiator, the probability of success of a predecessor attack is minimized when N, the number of peers in the first generation is minimized. Thus, examining equation (5) and recalling that we have fixed the load cost to C, the probabilistic anonymity is maximized when $f_N = 1$ and $E(\xi) = 1 - \frac{1}{C-1}$.

4.3. Network Operator and Colluding Peers

Recall that the anonymity set size, in terms of the network operator, is maximized when the expected value of N is maximized and that the probability of success is minimized as the expected value of ξ approaches 1. Thus, for a fixed load cost, their exists a trade off between maximizing the anonymity set size and minimizing the success of a predecessor attack. This trade-off is relevant both when the network operator and colluding peers act individually as discussed above and when the adversary consist of the network operator colluding with a set of peers, as such an adversary can launch a predecessor attack within the anonymity set.

Given this trade-off the optimal distributions $\hat{f}_\xi$ and $\hat{f}_N$ that maximize anonymity by balancing the anonymity set size and the probability of a successful predecessor attack, if they exist, are not immediately clear. A closer look into the optimal distributions $\hat{f}_\xi$ and $\hat{f}_N$ is reserved for future work.

References

[1] C. A. Ardagna, S. Jajodia, P. Samarati, and A. Stavrou. Providing mobile users' anonymity in hybrid networks. In *Computer Security–ESORICS 2010*, pages 540–557. Springer, 2010.

[2] F. Galton and H. Watson. On the probability of the extinction of families. In *Mathematical Demography*, pages 399–406. Springer, 1977.

[3] M. K. Reiter and A. D. Rubin. Crowds: Anonymity for web transactions. *ACM Transactions on Information and System Security (TISSEC)*, 1(1):66–92, 1998.