Herd: A Scalable, Traffic Analysis Resistant Anonymity Network for VoIP Systems

Stefan Savage
University of California, San Diego
savage@cs.ucsd.edu

It has long been understood that anonymity and efficiency are conflicting design criteria in the design of communication systems. Strongly anonymous systems such as those using Chaum's Mix networks can impose significant bandwidth and latency overheads, while other systems typically attain better performance at the expense of relaxing their resistance to traffic analysis. This paper describes a system called Herd that tries to thread this needle in the specific context of voice communication. While voice has stringent latency requirements, its call semantics and low constant bandwidth profile can significantly reduce the overhead of chaff traffic. Moreover, the authors of Herd explore how clever engineering can shift these overheads in a way that favors their application.

In many respects Herd is a traditional anonymity network, with hop-by-hop routing and layered encryption (like Tor) and chaff cover traffic (like Tarzan or P^5). However, it has a number of important innovations as well. First, communications is via dedicated infrastructure mix networks associated with individual "trust zones" run by distinct providers. In the proposed deployment, each trust zone would correspond to an individual provider operating under a single state jurisdiction (i.e., not across political boundaries). In turn, each zone operates a set of co-located mixes among which users can choose at random. This structure allows users to reason qualitatively about both commercial and state-level trust issues (in their choice of zone) and then, within a trusted zone, benefit from the guarantees provided by mix nets. Since each party can select independently, this may involve multiple zones, each with independent mixes, but since this is only a small number (1 - 3) of long distance links, latency is kept low. The other interesting technical contribution is that their design can optionally incorporate untrusted superpeers to offload CPU and bandwidth demands on infrastructure mixes. Superpeers offer transit bandwidth between clients and mixes, who use network coding to conceal which clients are in active conversations. As a result, the overhead on mixes can be limited to the number of active callers instead of the number of total participants.

The paper describes a working, open source, system, performs traffic analysis attacks on itself and competing systems, evaluates scalability and overhead using a real-life cell call database and performs a related comparative analysis of call quality offered by different systems.

This paper generated vigorous discussion at the PC meeting revolving primarily around two points: the value of superpeers and the completeness of security. The first issue reflects the fact that there is no free lunch. Superpeers do not reduce the overhead of the anonymity network – the costs must still be paid – but instead shift the resource demands from trusted to untrusted infrastructure. This is potentially an interesting tradeoff because untrusted resources may have different economic constraints than trusted resources (although this is an unsettled question and one worthy of further consideration). The second issue is the classic bugaboo of papers that propose new security systems – "how do we know they got it all right?" In some sense, this question is well formed since our history is littered with systems and protocols claimed secure that didn't anticipate an attack that subsequently rendered them vulnerable. While the authors make strong arguments for their security, Herd is a complex system and it is difficult to feel confident than all possible avenues for attack have been addressed. However, this problem is not unique to Herd and indeed, it is precisely via the publication of such efforts and their ability to withstand (or not) subsequent attack that we come to better understand these issues.

In summary, this paper describes and evaluates a real system that makes interesting tradeoffs to provide strong, but low-latency, guarantees for VOIP communication. This is a particularly opportune time for such research and we expect to see more of such work in the future.