Rangzen: Anonymously Getting The Word Out In a Blackout

Giulia Fanti
UC Berkeley
fanti@eecs.berkeley.edu

Adam Lerner
University of Washington
lerner@washington.edu

Yahel Ben-David
UC Berkeley/De Novo
yahel@eecs.berkeley.edu

Barath Raghavan
ICSI
barath@icsi.berkeley.edu

ABSTRACT

Recent years have seen the rise of a new type of state-level adversary. Governments have shown themselves willing to both surveil their citizens using Internet infrastructure and impose blackouts to shut off that very infrastructure during times of civil strife. However, it is exactly during such strife that citizens need secure, reliable, and anonymous communications the most.

In this paper, we present Rangzen[1], a system for anonymous broadcast messaging during Internet/cellular network blackouts. Rangzen is unique in both aim and design. Our aim is to provide an anonymous, one-to-many messaging layer that requires only users' smartphones and can withstand network-level attacks. Our design is a delay-tolerant mesh network which deprioritizes adversarial messages by means of a private social graph. We present extensive analytical and simulation results demonstrating Rangzen's efficacy and introduce what we believe is a fundamental trade-off between author anonymity in broadcast networks and the ability of a prioritization system to filter out undesirable content.

1. INTRODUCTION

Over the past decade, the balance of power between citizens and governments has tilted inexorably in the latter's direction. Notwithstanding old lore that the Internet was once the domain of pioneers and activists and a realm of free communications, it is increasingly coming under centralized control [61]. While this has proved to be a boon to the average user under placid conditions—large companies are able to deliver efficient and reliable services and governments are able to police networks for criminal activity—the opposite is true in times of unrest.

Indeed, during societal unrest, centralized infrastructure can be easily co-opted. In recent years, authorities in Egypt, Syria, and Iran, among others, have shut down their already heavily-surveilled Internet access during times when citizens were questioning those very authorities' political legitimacy [16, 23, 60]. However, it is exactly in such moments that citizens need the ability to communicate without restriction and risk of retribution.

In this paper, we present Rangzen, a system for anonymously spreading information at a large scale during a communications blackout. Two main challenges characterize our scenario: a) a lack of functioning infrastructure and b) powerful adversaries intent upon widespread interference and monitoring. We address the first challenge by broadcasting messages over a delay-tolerant, smartphone-based, mobile mesh network. This architecture enables reliable (albeit high-latency) communication. We address the second challenge with a social trust-based prioritization algorithm. Counterintuitively, Rangzen's prioritization algorithm allows users to receive trustworthy messages without knowing their origin or author; this property preserves authors' anonymity.

More concretely, our solution leverages the real-world properties of social trust and human mobility, each of which is a graph over individuals. The social trust graph is built by users via pairwise user trust establishments, while the physical proximity graph is implicit in the users' time-varying locations. Given this foundation, during message propagation we attempt to separate the wheat from the chaff. Such a task is difficult in any decentralized, anonymous system—adversaries can flood the network with bogus messages, and the lack of point-to-point communication makes organizing content difficult. We address this by using social trust information in a protocol for privately computing the size of a set intersection to yield a per-message trust metric. While our discussion focuses on an instantiation of this approach—Rangzen—our aim in this paper is to understand whether this system, or any system of its kind, could endure in a real blackout.

We evaluate the Rangzen approach through a combination of simulation and analysis. Our key aims are to a) understand how well a distributed, low-connectivity anonymity system can prioritize and deliver messages while withstanding attacks, b) quantify the anonymity levels possible in such a system, and c) explore the trade-off between anonymity and network prioritization efficacy. In the absence of ethnographic data from recent social unrest, we perform these evaluations over what we believe to be reflective social trust and physical proximity graphs. In simulation, we find that Rangzen is able to deliver a message to 80 percent of a storage-limited population within 20 hours of the ideal algorithm—epidemic routing over *unlimited*-storage devices. On average, 90 percent of simulated Rangzen nodes receive a legitimate message within 24 hours; this takes 50 hours on average for messages originating from adversaries. We also find that under certain mobility assumptions, if the adversary wishes to identify a set of nodes that contains the true author with probability greater than 0.8, that set should be at least as large as four-tenths of the network.

2. BACKGROUND

The design of a communication system for use in a government-imposed Internet blackout is a complex undertaking because of the socio-political dynamics that define both user desires and government actions. This understanding is crucial to building a system that is realistic in its aims and limits. In this section we present an abridged analysis of relevant socio-political turmoil and its implications for Rangzen.

2.1 Technology's Role in Revolution

Recent years have seen both successful and failed political revolutions in Tunisia, Libya, Egypt, Syria, Iran, Thailand, and

[1]Rangzen is the Tibetan word for freedom or liberty.

Ukraine [7, 8, 28, 68], and large-scale social demonstrations in Greece, Turkey, Spain, Canada, and the United States [9, 12]. In most (if not all), smartphone- and Internet-based organizing proved prevalent, and was widespread across social demographics. Perhaps not surprisingly, governments responded in kind, leveraging the same communications networks not only to surveil [30] but also to communicate directly with the population [66].

Implication: Smartphones are now sufficiently widespread that they can serve as democratic means for communication, though the communication they enable must be secured.

2.2 The Need for Anonymity

Technology's role in the spread of dissent is somewhat ambiguous. Although it certainly enables the spread of information, in many cases that very information has helped governments to target activists to an unprecedented degree. This targeting ranges from identifying dissenters online to harassing citizens who possess out-of-the-ordinary communication hardware [29].

Implication: During periods of social unrest, communication systems should protect individuals from being directly linked to objectionable content.

2.3 The Need for Robustness

The communications blackouts we consider in this paper are not accidents, and are due to a desire to stem the tide of public discontent. Nevertheless, the means by which blackouts have been implemented and their degree of totality have varied considerably. Some blackouts have involved BGP route withdrawals [19], while others have been more severe cuts [62]. In addition, total blackouts are made easier for governments to impose by a non-diverse network infrastructure with few providers and/or heavy government control. During these blackouts, governments are likely to try to thwart any temporary workarounds used by the population.

Implication: The cause or mechanism of a blackout should not impact the subsequent operation of the system, and the system should resist unilateral shut-down or co-option.

2.4 The Paradox of Fast Communication

Recently, political theorists have studied the impermanence of so-called "Twitter revolutions", commenting on the absence of slow, steady community organizing and in-person contact that has led to lasting political movements in the past [63]. Indeed, movements over the last century have often built momentum over decades, only to come to a sudden point of crisis involving drastic government actions such as the imposition of martial law or communications blackouts. The determining factor in the outcomes of such crises is largely an enigma, though the strength of the underlying movements has been a crucial factor [37]. While only history will adjudicate the impact of modern communications tools in political organizing, we believe the evidence suggests that rapid communication is not crucial.

Implication: The system need not enable rapid communication, but it should enable trustworthy and reliable communication.

2.5 Coalitions

The most recent uprisings in Egypt and Syria have clearly shown that dualistic thinking—government adversary versus citizen activist—is shortsighted. Rather, populations often factionalize into groups defined by political, religious, ethnic, and geographic characteristics; each group may see the others as adversarial [6].

Implication: The system must continue to function despite shifting coalitions of users and a dynamic notion of adversarial interaction.

		Communications	
		Point-to-point	**Broadcast**
Infrastructure	**Mesh**	Threshold-pivot [39], ALAR [45], Firechat* [52]	Rangzen
	Dedicated	Tor [25], VPNs/proxies [1, 2], Crowds [59], Tarzan [32], Free Haven [24]	Twitter* [5], Secret [3], DC Nets [14], Dissent [18]

Table 1: Design space of anonymity systems, with canonical examples. Internet-based systems are considered to use dedicated infrastructure. *Twitter and Firechat at best provide pseudonymity, but due to phone linkage often not even that.

3. DESIGN SPACE AND GOALS

While Rangzen belongs to a long lineage of anonymity systems, we believe that both its target environment and approach are unique in the literature. In this section, we briefly consider prior work and then describe Rangzen's threat model and goals.

3.1 Related Work

Table 1 compares Rangzen with a number of other related anonymity systems. We believe Rangzen is unique in its support of anonymity in a mesh-based, one-to-many communication paradigm. Rangzen emerges from the convergence of several research areas. Our work relates to Sybil defense [56, 65, 67, 69] through our use of social graph structure to distinguish users. Similarly, there is a rich literature on private set intersection over social data to achieve privacy and anonymity [43, 44]. We build thematically on other systems involving distributed, privacy-preserving trust and reputation [33, 38] and, in addition to those systems in Table 1, owe a debt to previous work on information leakage and hiding in social networks (e.g. [22]).

3.2 Threat Model

An attacker's objectives in our setting could include preventing the spread of legitimate messages, spreading adversarial messages, and deanonymizing users. Deanonymization in this context could include one or more of the following: discovering the author of a message, matching real-world people to Rangzen nodes, and learning the existence of the social/trust links between (anonymous) nodes in the system. We describe attacks aimed at these goals.

Physical/MAC Layer Attacks. An attacker might use physical or MAC layer attacks to disable nearby nodes and prevent them from communicating. In a wireless setting, this would require the attacker to deploy many devices for coverage. We assume the attacker is resource-limited in deploying such devices.

Device Capture. A primary device capture attack involves the adversary taking physical control of a legitimate user's smartphone and thereafter pretending to be the user in all interactions (except further social trust link establishment). A secondary device capture attack is one in which the adversary deploys targeted malware that exploits the smartphones of legitimate users.

Device Impersonation. In a device impersonation attack, the adversary learns some identifying information of a user or set of users without device capture and then attempts to impersonate those users' devices in peer interactions.

Ordinary Participation. In an ordinary participation attack, the adversary deploys devices which implement the Rangzen protocol, possibly in a Byzantine, malicious fashion. These nodes take part in

		Impact		
		User Identification	Denial of Service	Propaganda
Attack	Physical/MAC Attacks		●	
	Device Capture	●	○	●
	Device Impersonation		○	●
	Ordinary Participation	○	○	○
	Message Spread	○	○	●
	Trust Graph Extraction	●		

Table 2: Taxonomy of attacks and their potential impact on Rangzen. ● indicates when the attack directly enables the impact indicated; ○ indicates when the attack only indirectly or partially enables the impact.

network activity—relaying messages, engaging in peer exchanges, etc.—while modifying the protocol in any way the attacker likes.

Message Spread. A message spread attack is an "ordinary participation" attack (see above) in which the attacker spreads its own messages. However, these messages, by virtue of their origin (an adversarial node), are undesirable and should not spread better than non-adversarial messages. When employing this attack, the adversary may create an arbitrary number of Sybil nodes [27], some of which may attempt to establish trust links.

Trust Graph Extraction. A trust graph extraction attack involves the adversary engaging in peer exchanges ("ordinary participation") with the intention of learning the underlying social trust graph. To do so, the adversary must establish real trust links with real users and then perform exchanges with users pretending to have only one trust link. (We expand upon this attack in later sections.)

Other Attacks. An attack on Rangzen's network health would have to either restrict human mobility—which is difficult at scale—or prevent phones from being able to communicate. If an adversary can prevent a nation of smartphones from communicating, it can likely thwart any mesh-based blackout solution; cell phones represent an upper bound on device mobility, which is strongly correlated with mesh network robustness [35, 53].

An out-of-band attack on Rangzen's prioritization scheme would involve manipulating the social graph. We address this by taking a cue from the sybil defense literature and assuming that adversaries can form only a bounded number of *attack edges* in the social graph [69]. That is, adversaries are limited in their ability to make friends with real users. Analogously to sybil defense systems, we leverage this limitation to deprioritize messages from attackers.

Author deanonymization could occur through a number of channels: social engineering, user-supplied message content[2], malware, or leaked algorithmic information (trust graph and computed message priority scores). In addition, we expect that the adversary monitors social relationships through outside means, and that this enables author deanonymization. Our analysis in §5 shows that the trust graph is well-protected and priority scores can be tuned for higher anonymity, but priority scores are Rangzen's weakest link against out-of-band attacks. As discussed in §5.4, this is inherent to trust-based prioritization systems.

Targeted attacks and non-goals. While an ideal system would be capable of defending against all the attacks described above, we consider an adversary that has massive technological resources but a limited number of agents in the population. We find such an adversary to be reflective of the scenarios described in §2. In addition, visual identification of a user, physical device identification via RF signatures, and other out of band attacks are relevant, but fall within our definition of targeted attacks, wherein the adversary must pursue a single individual; we do not believe that there are technological means of preventing such attacks.

[2]It is unlikely any system can defend against this.

Table 2 summarizes the impact of the listed attacks on Rangzen. We do not intend for Rangzen to perfectly thwart an adversary, but rather to enable anonymous citizen communication in the face of government oppression. Put another way, we aim for Rangzen not to prevent attacks, but to carry on despite them.

3.3 Natural Formation of the Problem

We believe Rangzen constitutes a natural approach to enabling anonymous communication during a network blackout. Put simply, communication in a blackout requires a non-infrastructure solution (a mesh of users communicating over ad-hoc links). Since low-bandwidth, mobile mesh networks are not good at real-time communication [36], delay-tolerance is necessary.

A natural metric in this setting is message propagation. While faster propagation is better, propagation speed is upper-bounded by the density and mobility of users. The desired speed of propagation is application-dependent. For Rangzen, distributing messages to a large fraction of users within an urban center with delays of tens of hours is sufficient to enable meaningful public discourse and organize events in time for participants to attend.

The ability of such a system to filter out undesirable messages (and consequently, deliver legitimate content faster) conflicts with message-authorship anonymity. We believe this trade-off is fundamental and unavoidable; message-authorship anonymity is the correct notion of anonymity in such a system because hiding protocol participation is hard [34].[3]

Finally, while Rangzen provides support for microblogging, the notion of broadcasted datagrams is quite general. As such, we suggest that this trade-off between anonymity and content prioritization applies to the entire space of blackout-resistant systems.

4. ARCHITECTURE AND DESIGN

The observations in §2 require Rangzen to resolve a fundamental tension between anonymity (which demands that users' identities be hidden) and robustness to network-level attacks (which requires some notion of reputation or identity). Many practical and academic anonymity systems attempt to resolve this tension using pseudonymity [4,39], but pseudonymous systems are susceptible to side-channel information correlation attacks [51, 58]. The problem is even more challenging during a blackout, which prevents the use of online trust mechanisms, such as Bitcoin-based protocols [49]. Rangzen addresses these challenges by providing anonymity rather than pseudonymity in message transmission and by leveraging a privacy-preserving social trust mechanism.

4.1 Design Overview

Rangzen forms a mobile ad-hoc network of smartphones. Users build a *trust graph* in which each user is a node, and graph edges

[3]This requires network participation to be legal or rarely prosecuted. Many blocked censorship circumvention tools remain widespread in dissent networking hotbeds today.

represent real-life trust relationships between people. Opportunistic encounters can occur between any two nodes, and data transfer volume depends on the encounter duration; thus, the system must decide which messages to transmit and how to order messages on the user's screen. Dedicated-infrastructure systems often choose this ranking based on message authorship and/or timestamps. To preserve anonymity, we instead rank messages by their trustworthiness, which is computed from the trust graph. The key intuition of our system is that **users assign greater trust to other users with whom they share many mutual friends**. Messages from untrusted users are the least likely to be displayed to the user, the last to be transmitted in an encounter, and the first to be discarded under storage pressure. In this section, we briefly describe the protocols and techniques we use in Rangzen to establish the trust graph, communicate between users, generate messages, and prioritize resources.[4]

4.2 Trust Graph Setup

Here we describe the logistics of building and maintaining a trust graph. Borrowing from the literature on social graphs, we refer to a pair of nodes that trust one another as "friends".

4.2.1 Trust Establishment

To make it difficult for adversaries and their sybils to establish genuine trust links with ordinary users, our trust establishment protocol requires the face-to-face exchange of information that is visually present only on users' devices. Each user has a (time-varying) opaque pseudonym, or ID. These pseudonyms are not attached to messages—messages remain entirely anonymous, with these IDs used only for social-network based message prioritization. When two nodes confirm mutual trust, each transmits her user ID and a hash of her name to the other node. This adds an edge between the two nodes in the trust graph. Thus each node stores a list of her friends' IDs and hashed names. Nodes cannot add or remove elements from this list without entering a password, and deniable encryption allows users to decrypt the hashed name list to a semantically indistinguishable list of hashes under duress [13]. Note that adding edges to the trust graph is completely distributed, so there is no global record of the social graph. When two users establish a trust relationship, they exchange public keys via a public-key exchange protocol that ensures that each user has verified the physical presence of the other [31]. Specifically, each party verifies that the other's public key is authentic via challenge-response.

4.3 Peer Encounter Protocol

Rangzen's mesh communication is achieved via lightweight, pairwise message exchanges. Here we overview the protocol enacted by two peers when they encounter one another, including both the exchange of information and how each node quantifies social trust and uses it for message discrimination.

4.3.1 Lightweight Protocol

In our prototype Android implementation (not considered in detail in this paper), we leverage in-built support for WiFi-direct, which enables rapid association-free communication between smartphones. In preliminary tests, we found that the devices could reliably connect to a peer and transmit 2.5 MB of data within 30 seconds when the devices were up to 30 meters from one another. In the Rangzen protocol, one full peer exchange can complete in 3 one-way UDP packet transmissions, enabling lightweight and rapid opportunistic message exchange.

[4]However, our evaluation does not depend upon protocol specifics but rather is general to systems that use the Rangzen approach.

4.3.2 Social Trust Metric

Our basic social assumptions are that a) humans are more likely to want to see messages from people with whom they share many mutual friends, and b) adversarial nodes are unlikely to share many mutual friends with anyone. To capture these ideas, each pair of nodes computes a social trust score during an opportunistic encounter. This trust score $T(\boldsymbol{a}, \boldsymbol{b})$ describes how much $\boldsymbol{a}$ trusts $\boldsymbol{b}$. In words, this asymmetric function is the ratio of mutual friends between $\boldsymbol{a}$ and $\boldsymbol{b}$ over the total friends of $\boldsymbol{a}$:

$$T(\boldsymbol{a}, \boldsymbol{b}) = \max\left(\frac{F(\boldsymbol{a}) \cap F(\boldsymbol{b})}{F(\boldsymbol{a})}, \epsilon\right)$$

where $F(\boldsymbol{a})$ denotes the set of $\boldsymbol{a}$'s friends, and ϵ is a small positive constant that ensures that ordering is preserved—even if the nodes share no mutual friends. To compute the numerator of $T(\boldsymbol{a}, \boldsymbol{b})$ in a privacy-preserving manner, we leverage existing cryptographic techniques for private set intersection with cardinality (PSI-Ca) [20]. This lets two nodes compute how many mutual friends they share, without ever learning *which* mutual friends they share, or any information about non-shared friends. [20] is secure only against semi-honest adversaries. There also exist schemes secure against malicious adversaries [26, 50], but for the sake of protocol efficiency, we use [20], and argue in Section 6.2 that there is no incentive for an adversary to misbehave during the PSI-Ca protocol. Three one-way transmissions are required. We estimate that for text messages, three transmissions will suffice for transmitting hundreds of messages along with the private set intersection data.

4.3.3 Mapping from trust to priority

Nodes use the trust function $T(\boldsymbol{a}, \boldsymbol{b})$ to prioritize messages arriving from another node. Each message in Rangzen is associated with a priority score in $[0, 1]$. If $\boldsymbol{a}$ receives a message from $\boldsymbol{b}$ with priority p_o, then $\boldsymbol{a}$ will insert the message into her queue with a priority that is a sigmoidal function of the trust score:

$$Tr_0^1[(p_{b,a}(T(\boldsymbol{a}, \boldsymbol{b})) \times p_o) + z_a],$$

where $z_a \sim \mathcal{N}(\mu, \sigma^2)$ is additive Gaussian noise used to improve message propagation,[5] $Tr_0^1[x]$ is a threshold forcing x to be in the range $[0, 1]$, and

$$p_{b,a}(T(\boldsymbol{a}, \boldsymbol{b})) = \frac{1}{1 + \exp\{-\rho(T(\boldsymbol{a}, \boldsymbol{b}) - \tau)\}}. \quad (1)$$

Approximately, this means $\boldsymbol{a}$ will trust $\boldsymbol{b}$ fully if the ratio of mutual friends is greater than τ. We used $\rho = 13$ and $\tau = 0.3$ for a sigmoid that transitioned sharply in the range $[0, 1]$. These constants will need to be tuned in system tests. If a device runs out of storage, the lowest-priority messages get dropped first.

4.4 Message Management

Rangzen enables users to identify messages that are likely to be relevant—messages with high trust scores, indicating that they were originated and/or propagated by socially proximate nodes.

4.4.1 Authoring

Users are presented with an ordinary text-message-like message interface, but without from/to fields. New messages are initialized to have priority 1. If a user particularly likes a message from another node, she can manually "upvote" the message, which resets

[5]This noise parameter helps unpopular nodes spread content by randomly increasing (or decreasing) priority scores, but it also improves author anonymity. We demonstrate these claims in §5.

its priority score to 1. Note that since messages are anonymous, upvoting a message is equivalent to reauthoring it.[6]

4.4.2 *Organizing and Prioritization*

We leverage the social trust metric to prioritize messages in the display presented to the user. given topic tag. The highest-ranked messages are also the first to be transmitted during an opportunistic encounter and the last to be discarded under storage pressure. To prevent old messages from pushing out newer messages, each clients degrade a message's priority as a function of time. This happens via random (in time and magnitude) priority degradations to avoid storing timestamps.

5. EVALUATION

In this section we analyze not only Rangzen, but also Rangzen-like systems that employ a delay-tolerant mesh to enable anonymous broadcast communication, to better understand how well such systems can circumvent large-scale Internet blackouts.

5.1 Evaluation Setup

We evaluated Rangzen's message propagation and anonymity properties using both simulation and analytical techniques. Real-world datasets informed our assumptions as much as possible, including mobility traces (EPFL Cabspotting [55], St. Andrews Locshare [11], University of Milano PMTR [48], and Technicolor SIGCOMM [54]) and social graphs (two subgraphs of the Facebook social graph [47, 64]). We also tested our algorithms on datasets of mobility *and* social connections [11, 15] (in principle, this is what we want), but we found the datasets to be too sparse in time (Gowalla) and space (St. Andrews) for effective evaluation.

Our simulator consists of 2100 lines of Java code built upon MASON [46], a discrete-event multiagent simulation library. Our simulator accepts social network data inputs or can generate scale-free random social graphs when absent in a dataset [10]. The simulator supports various mobility datasets. It replays agent locations over time and agents within 20 m[7] are made to encounter one another with some small probability (we chose 0.05). This is meant to simulate unreliable message exchanges for worst-case evaluations.

Nodes can also be specified as adversarial nodes which perform physical/MAC layer attacks. We model these attacks in a worst-case analysis by assuming that all nodes within range of the attacking node are unable to communicate at all.

5.2 Message Propagation

Our metrics of success for message propagation are a) the time required for a tagged message to reach 90 percent of the honest population, and b) the fraction of honest nodes that have received a message by a given time. These metrics are chosen for use cases like protest organization, in which mobilization is only possible if a large portion of the population acts in cooperation. All plots are averaged over 40 runs. We consider epidemic propagation over infinite-storage devices an upper bound on the spread rate since a mesh DTN cannot disseminate content faster than message flooding if storage is unconstrained. We use the Cabspotting mobility dataset [55] with a randomly-generated Albert-Barabàsi social graph [10]. The Albert-Barabàsi generative model lets us create arbitrarily-sized social networks, and it displays some common properties of social networks, like high clustering-coefficient, power-law degree distribution, and short path lengths between nodes. However, it does not capture some other properties of social graphs, such as community development. Also, true social graphs are typically correlated with mobility patterns.

[6]It is imperative that users not inadvertently reveal their identities through message content; unfortunately, this is very difficult to prevent with technological solutions alone. As with all anonymous communication systems, user education is critical and interfaces must be designed in ways that make doing the safe thing easy.

[7]Our physical tests found reliable communication possible at ranges up to 30 m.

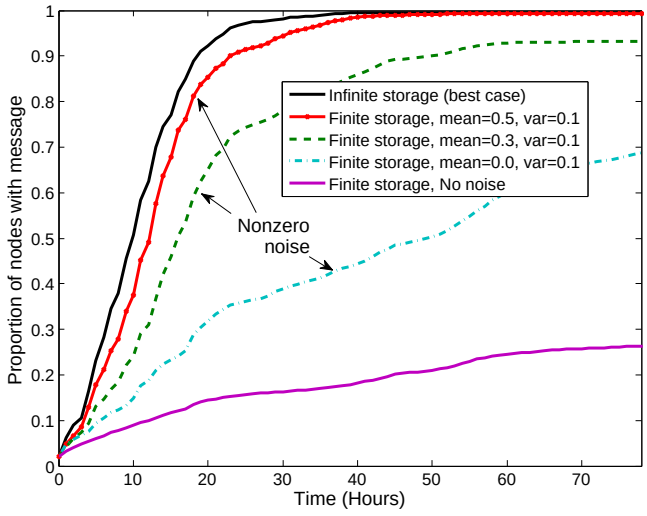

Figure 1: Impact of Rangzen protocol on legitimate message propagation without an adversary. The additive Gaussian noise in priority scores clearly improves propagation, but may hamper the system's ability to filter out adversarial messages.

5.2.1 *Propagation without an adversary*

Figure 1 shows the propagation of legitimate messages with no adversary. The curves represent different distributions of the noise parameter z_i in our trust metric. Figure 1 suggests that even using random social graphs, **Rangzen can reach at least 80 percent of the population within 24 hours and 90 percent of the population 20 hours after infinite-storage epidemic routing does so**. Since taxis move city-wide (unlike most citizens), this experiment is most representative of the time it takes a message to span a city.

5.2.2 *Propagation with a passive adversary*

Next, we demonstrate the performance of Rangzen under a passive adversary, which deploys devices that follow the Rangzen protocol, but may also disseminate their own content. This adversary is related to the coalitions mentioned in §2.5. Coalitions, or even distinct groups of friends, may wish to emphasize their own content internally without directly attacking other coalitions' communications. We expect nodes that do not belong to coalition C to have few friends in C, and therefore low degree in C's subgraph. Passively adversarial messages can thus be thought of as messages from an unpopular node. (We also expect our global adversary to be unpopular in the global trust graph.) Figure 2 shows the effects of node popularity on propagation speed. Here, (un)popular nodes were selected randomly from the 5 percent worst- or best-connected nodes in the social network. The figure shows that well-connected nodes have a distinct advantage; **messages from popular nodes reach 90 percent of the population as much as 40 hours earlier than messages from unpopular nodes, for certain noise levels**. This model of communication is not necessarily optimal, but it is consistent with human communication patterns; in the real world, unpopular people cannot reach a broad audience without the help of a popular person. In Rangzen, this corresponds to a

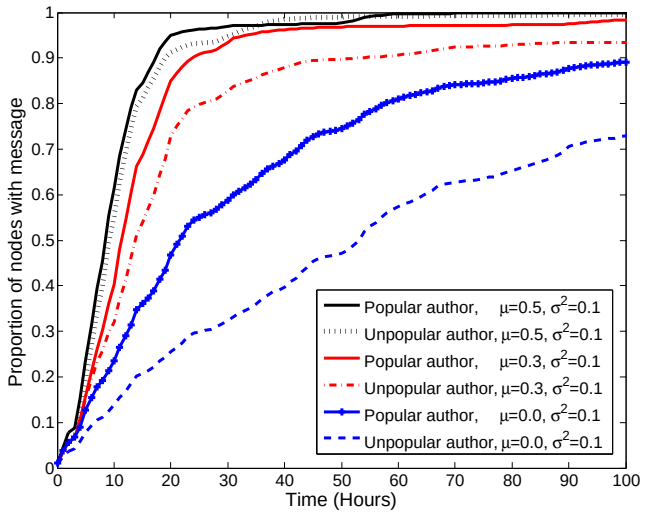

Figure 2: Popular nodes can spread messages faster than unpopular nodes. This effect is more pronounced when nodes add less noise prior to transfers (e.g. lower μ). We expect adversarial nodes to be unpopular.

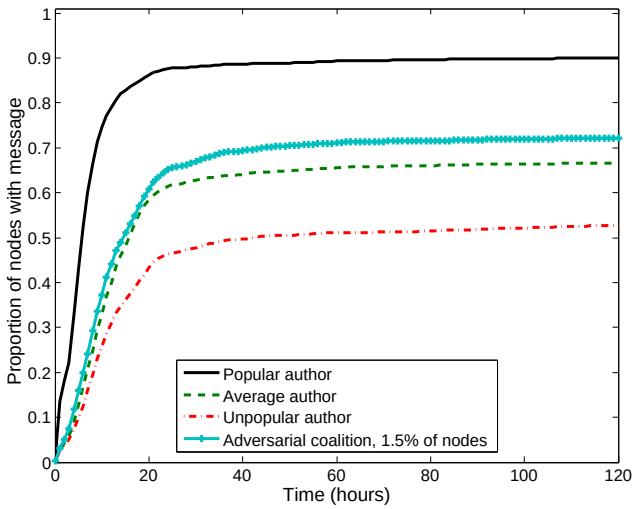

Figure 3: Ability of global adversaries to spread propaganda.

central node "upvoting" a message.

5.2.3 *Propaganda Message Spread*

We now consider an active adversary that controls a small fraction of nodes—this represents an active, global adversary. Adversaries individually have few friends, but they increase their effective degrees by sharing friend IDs. This coalition spreads only its own messages. It can create Sybils, but since Sybils cannot befriend honest nodes more than human adversaries, we find there is no advantage to creating Sybils. We used noise parameters of $\mu = 0.0$ and $\sigma^2 = 0.1$ with a community of 400 nodes, which utilizes almost the entire Cabspotting dataset.

Figure 3 illustrates the propagation time of messages originating from popular, average, and unpopular honest nodes, as well as the adversarial coalition; we assume that 1.5 percent of the population belongs to the adversarial coalition.[8] The figure shows that **the adversarial coalition can spread messages a little bit better than average nodes, but at least 30 percent worse than *individual* popular nodes**. This happens because the coalition of malicious nodes continuously broadcasts high-priority propaganda, while an average node has no such group to help with propagation. Therefore, as long as the adversary has limited friends within a coalition, popular nodes can use Rangzen to spread information reliably to the masses. At very small scales (50 nodes), we observed that average and unpopular nodes actually performed *better* than the adversarial coalition for the first 48 hours. This suggests that the app can still be used within tighter social circles to communicate, though one should be well-connected to communicate at a large scale.

5.2.4 *Robustness to physical/MAC attacks*

Modeling physical/MAC-layer attacks is difficult because the adversary's intent and tools can vary widely. However, we wish to show that the adversary cannot prevent communication of Rangzen users across an entire city. For a worst-case estimate of the effect of physical or MAC-layer attacks on message propagation, we first consider a physical-layer attacker (such as a jammer). We model the jammer as a point source of radiation in one of the WiFi frequency bands (20 MHz bands at either 2.4 GHz or 5 GHz), as a best-case for the attacker. The power of the radiation received by the client is assumed to follow the path loss formula:

$$P_R = P_T \left(\frac{c}{4\pi df} \right)^2,$$

where c is the speed of light, f is the signal frequency in Hz, d is the distance traversed, and P_T and P_R are the transmitted and received power, respectively (we assume equal antenna gains). We ignore factors like reflection, diffraction, and absorption, which would significantly weaken a jamming adversary, especially in a city setting. We estimate the transmit power of a smartphone to be 251 mW (corresponding to average output power over the 5.4 GHz band), and we estimate the maximum output power of a stationary jammer to be 20 W in the same WiFi band. We chose this number based on a search of commercially available jammers. Although many products advertise high power output, the bandwidth of such products is typically high. Under these assumptions, a stationary jammer would need to be within roughly 180 m of the receiver (with a line-of-sight connection) to plausibly jam a transmission happening between two nodes 20 m apart. A mobile jammer would have to be even closer due to power constraints.

However, is is possible that by using MAC-layer attack techniques, the attacker could extend the range of the attack beyond hundreds of meters—for instance, by sending lower-power messages that cause other nodes to not transmit. We allow for this, and in Figure 4 show the impact of geography-based attacks (such as common attacks on the physical or MAC-layer) on message propagation. For a worst-case estimate, we consider these lower-layer attacks at ranges well beyond the 180 m distance that would be a limit of high-end commercial jammers in WiFi frequencies.

Figure 4 considers both mobile and stationary adversaries that are either incidental (non-optimally placed for human mobility patterns) or adversarial (optimally placed). We model the mobile, incidental adversary as one of the nodes in the mobility trace. Similarly, the incidental stationary attackers are placed uniformly at random within the simulation area.[9] To choose an approximately-optimal placement for an adversarial, stationary attacker, we use a simu-

[8] At its height, the Stasi employed 0.6% of the East German population as agents and 1.5% of the population were either employed or in collaboration [41].

[9] This might correspond to police stations distributed in a city.

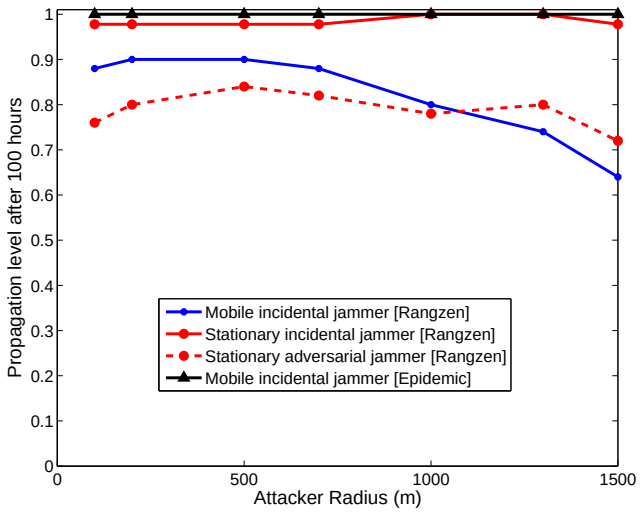

Figure 4: Impact of physical/MAC attacks on propagation.

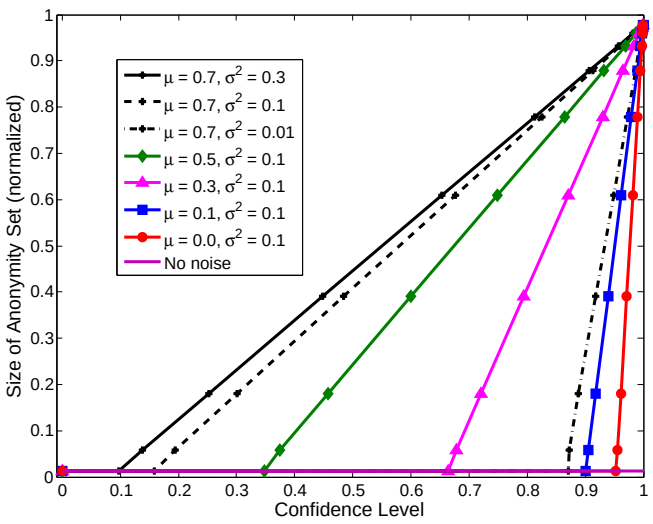

Figure 5: Author anonymity set size (fraction of population) as a function of the estimator's confidence level, given a node stores the message with priority 1. The point $(0.5, 0.2)$ means that the smallest set of nodes to include the author with probability 0.5 contains at least 20 percent of the network.

lated annealing algorithm for base station placement in DTNs [40]. An optimal adversarially-selected *mobile* attacker would have to know the entire population's location at every instant in time, and then solve an NP-hard problem [40]. Consequently, we don't believe that optimally mobile adversaries pose an elevated risk over nodes that traverse popular routes regularly (e.g. taxi cabs).

Figure 4 shows that **even when physical/MAC-layer attackers have ranges as high as 1000 m, the system propagates at least 80 percent as well as in a best-case scenario**. Such an attack is unlikely in practice, but this highlights Rangzen's robustness to geographically-localized attacks.

5.3 Privacy Guarantees

In this section we evaluate Rangzen's anonymity properties and its resistance to an attacker who wishes to pinpoint a message's author or extract information about the social trust graph.

5.3.1 Authorship deniability

Claim: Sending a high-priority message doesn't necessarily make you look like the author.

If the adversary receives a high-priority message from an honest node, Rangzen should enable the sender to plausibly deny authorship. To capture this, we measure the size of the *anonymity set*, or the set of nodes that could have plausibly authored a particular message. For high deniability protection, the anonymity set should be as large as possible. To quantify the set size, we estimate how many hops a given message took in the (time-varying) connectivity graph, and then estimate how many nodes are that many hops away.

More precisely, we compute the probability mass function (pmf) of the number of hops a message has traversed before reaching a target node, given the observed priority score. Concretely, suppose node A receives a message from B. Let N denote the number of hops the message traversed *before* reaching B (e.g., if B were the author, $N = 0$). Let $S \in [0, 1]$ denote the priority with which A receives the message from B (before taking into account the mutual friends between A and B). Let Ω denote the event that the message is observed by a randomly-selected node (in this case, A). For a worst-case analysis, assume that A receives the message with priority $S = 1$. We wish to compute

$$P(N = n|\Omega, S = 1) = P(S = 1|N = n, \Omega)P(N = n|\Omega).$$

We describe our modeling of $P(N = n|\Omega)$ and $P(S = 1|\Omega, N = n)$ in Appendix B. Given these models, we can estimate $P(N = n|\Omega, S = 1)$ as a function of n. Combining this with a mobility dataset, we can numerically estimate the size of the anonymity set for a given trace. Figure 5 illustrates the size of the author's anonymity set as a function of the estimator's confidence level (i.e. the probability that the true author is in the anonymity set) for the SIGCOMM social graph and dataset [54]. Using $\mu = 0.3$ and $\sigma^2 = 0.1$, the set of nodes that contains the true author with probability 0.9 contains 80 percent of network nodes. More noise significantly increases the anonymity set size, leading to greater authorship deniability. The correct noise parameters should be selected empirically to balance anonymity with message propagation.

5.3.2 Unlinkability of nodes to identities

Claim: The amount of information the adversary can learn about the trust graph is limited.

Here we consider an adversary who aims to learn global information about the trust graph of the Rangzen network, such as which pairs of nodes are friends. This information is dangerous because of the risk of deanonymization—the likelihood that the graph can be correlated with social graphs in other networks (e.g., Facebook, Twitter) to learn the real identities of users [51].

Our goal is to quantify how much information the adversary learns about the true social graph through attacks on the private set intersection protocol. Throughout, we will assume the adversary knows the vertex set V of the social graph, since use of Rangzen is not assumed to be inherently incriminating. There are many definitions in the literature for graph information content (see [21] for a review). None of the definitions is clearly superior, so we use the proportion of common edges as a heuristic metric. That is, if the original graph is denoted $G = (V, E)$ and the subgraph is denoted $G_s = (V, E_s)$ with $E_s \subseteq E$, then our similarity metric is

$$d_\epsilon(G, G_s) = \frac{|E_s|}{|E|}$$

This metric is closely related to the definition of graph entropy by Rashevsky et al [57] and was also shown to be strongly correlated

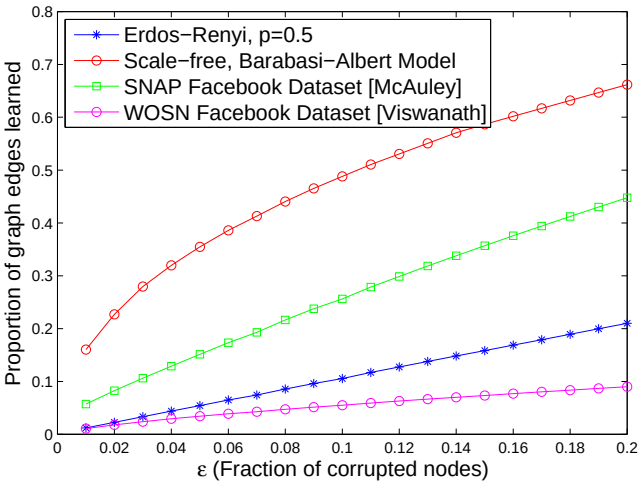

Figure 6: Proportion of graph edges learned by the adversary (d_ϵ) as a function of ϵ (proportion of corrupted nodes).

with deanonymization success in [51]. Assuming the adversary can corrupt at most fraction ϵ of the nodes, we wish to upper bound d_ϵ as a function of ϵ. This section demonstrates that the adversary will be unable to learn more than 15 percent of the graph edges by corrupting up to 5 percent of the nodes, and this quantity can be further limited by artificially adding and removing edges from the social graph during PSI-Ca interactions.

Static graph. We start by assuming the trust graph does not change. As time tends to infinity, we assume that the adversary can learn all edges emanating from nodes corrupted by the adversary. This is a worst-case estimate, because it assumes that the adversary knows how to align its learned subgraph within the larger trust graph (or a similar social graph from a different domain). In practice, subgraph alignment is not trivial.

Figure 6 illustrates the proportion of edges learned as a function of the proportion of nodes corrupted. The SNAP dataset is a Facebook ego-social-circle dataset [47], and the WOSN dataset contains social connections between 55,000 nodes in the Facebook New Orleans network as of 2009 [64]. The figure suggests that as long as the adversary cannot corrupt more than about 5 percent of nodes, it can learn at most 15 percent of the social graph. This estimate is worst-case; along with the subgraph alignment issues mentioned earlier, corrupting nodes is difficult, and we expect trust establishment to be less promiscuous in Rangzen than in Facebook.

Dynamic graph. Next, we assume that the graph is changing with time. This is a more realistic assumption, since people add and (less frequently) remove friends. In this case, we wish to ensure that the adversary's learning rate is dominated by the rate of change of the social graph. Such a system is difficult to characterize in general, so we will use a simple model to build intuition.

Consider three bins: one with edges that have been learned by the adversary (L), one with edges that have not been learned by the adversary (U), and a final bin containing edges that are not in the graph (X)—i.e., pairs of nodes that are not connected. Each time a new trust relationship is created in this trust subgraph, another edge is added to the U bin, and each time an edge is deleted (i.e. someone "unfriends" an acquaintance) an edge is removed from the L or the U bin. For a worst-case estimate of privacy, we will assume that the adversary knows when edges are deleted. Edges move from U to L whenever the adversary learns another edge in the graph. Thus we wish to characterize $|L|/|L+U|$. Recall

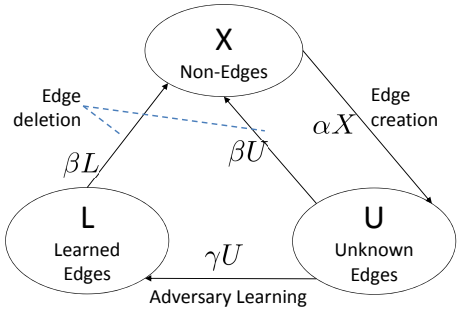

Figure 7: Adversarial learning of a dynamic trust graph.

that with a static graph, the adversary could learn at most a small fraction d_ϵ of the total edges in the graph. As such, our dynamic model operates within a restricted space of nodes and edges. For instance, if the adversary corrupts 5 percent of network nodes, then $N_E = X + L + U$ equals the number of edges possible between the corrupted 5 percent of nodes and the rest of the network. Any equilibrium value of d_ϵ in our dynamic model should therefore be multiplied by the results for the static graph (assuming people will never add false edges to the social graph).

Our underlying model for this system is a continuous-time Markov chain with Poisson events. Unfortunately, the state space of such a Markov chain grows exponentially in the number of total possible edges (N_E). As such, we can use a mean-field approximation, much like the analysis in [70]. Figure 7 illustrates our model of the system. αX is the rate of edge creation, $\beta(U+L)$ is the rate of edge deletion, and γU is the rate at which the adversary learns new edges.

We know that $X(t) = N_E - L(t) - U(t)$ where N_E describes the number of total possible edges. Letting $V(t) = [L(t)\ \ U(t)]^T$, we have a nonhomogeneous time-invariant linear system:

$$\frac{dV(t)}{dt} = \begin{bmatrix} -\beta & \gamma \\ -\alpha & -(\alpha+\beta+\gamma) \end{bmatrix} V(t) + \begin{bmatrix} 0 \\ \alpha N_E \end{bmatrix} \tag{2}$$

OBSERVATION 5.1. *Let $V(t) = [L(t)\ \ U(t)]^T$, with dynamics described in Equation 2. Then*

$$\lim_{t\to\infty} \frac{L(t)}{L(t)+U(t)} = \frac{\gamma}{\gamma+\beta}.$$

PROOF. (Sketch) It is straightforward to show that dynamical system (2) is internally stable, and the exact solution is

$$\begin{bmatrix} L(t) \\ U(t) \end{bmatrix} = \begin{bmatrix} \frac{\alpha\gamma N\left(\alpha-\gamma+(\beta+\gamma)e^{-(\alpha+\beta)t}-(\alpha+\beta)e^{-(\beta+\gamma)t}\right)}{(\alpha-\gamma)(\alpha+\beta)(\beta+\gamma)} \\ \frac{\alpha N\left(\beta(\alpha-\gamma)+-\alpha(\beta+\gamma)e^{-(\alpha+\beta)t}\gamma(\alpha+\beta)e^{-(\beta+\gamma)t}\right)}{(\alpha-\gamma)(\alpha+\beta)(\beta+\gamma)} \end{bmatrix} \tag{3}$$

We then consider

$$\frac{L(t)}{U(t)+L(t)}.$$

Since the exponential terms in (3) tend asymptotically to 0, the ratio of interest converges precisely to $\gamma/(\gamma+\beta)$. □

Figure 8 illustrates these analytic results compared to discrete stochastic simulated results. The colored bands are inter-quartile ranges over 40 trials. These results reflect only on the quality of our mean-field approximation, not on the assumption of constant-rate learning and social graph alterations. However, our model does capture the observation that social graph properties stabilize globally over time, despite continuing to fluctuate locally [42].

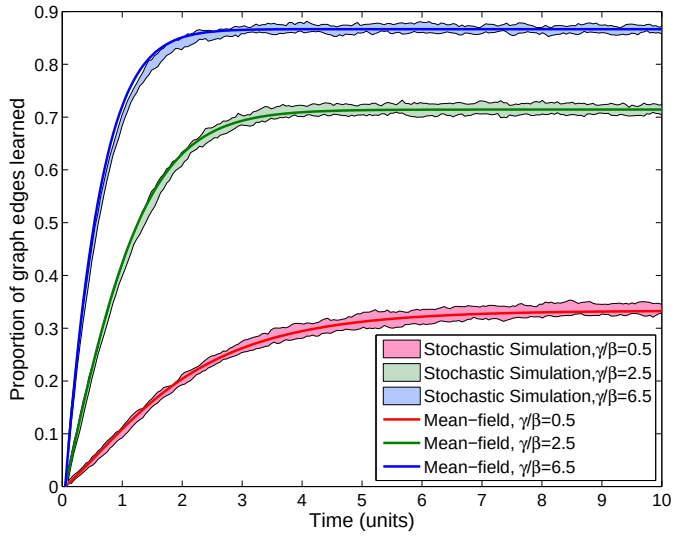

Figure 8: Adversarial graph learning over time, parameterized by the adversary's learning rate γ. Asymptotically, the leaked proportion of graph edges depends exclusively on the adversary's learning rate γ and the network-wide edge deletion rate β. We can therefore increase privacy by systematically omitting trust graph edges.

This result tells us two things: 1) if no edges are deleted, the adversary eventually learns the entire graph, and 2) asymptotic behavior is independent of the edge creation rate. Over a long time scale, we cannot rely on natural social graph growth to limit the adversary's knowledge. Instead, we should artificially simulate the deletion of edges by (for instance) including random subsets of users' friend sets in each private set intersection. These observations suggest that social graph-based deanonymization is a surmountable threat within our framework. However, tuning the relevant parameters will require more detailed field study over extended periods to better understand the trust graph's evolution.

5.4 Anonymity-Prioritization Tension

Claim: There is a trade-off between protecting the identity of message authors and meaningful message prioritization.

Our results in Figures 2 and 5 suggest a trade-off between message prioritization (i.e., the propagation differential between honest messages and propaganda) and anonymity (i.e., the size of the anonymity set). In this section, we show that for a simple problem, one-to-many communication presents a fundamental trade-off between author anonymity and the ability of the system to distribute messages to interested parties.

Suppose we start with a collection of nodes. One of these ($\boldsymbol{z}$) composes a message m. Nodes that have seen m are assigned value 1 (0 otherwise). The goal of the system is to spread the message only to nodes that want to see m. More precisely, let the function $d(\boldsymbol{x}, \boldsymbol{z}, m)$ denote the *semantic distance* between nodes $\boldsymbol{x}$ and $\boldsymbol{z}$ with regards to message m. Intuitively, $d(\boldsymbol{x}, \boldsymbol{z}, m)$ tells us the relevance of a message to a node; if $d(\boldsymbol{x}, \boldsymbol{z}, m)$ is very small, then m is very relevant to $\boldsymbol{x}$, and $\boldsymbol{x}$ should receive the message. Our primary modeling assumption is that the prioritization system reveals m to $\boldsymbol{x}$ with a probability that is a function of $d(\boldsymbol{x}, \boldsymbol{z}, m)$:

$$P(\boldsymbol{x} = 1 | Y = \boldsymbol{z}) = g(d(\boldsymbol{x}, \boldsymbol{z}, m)), \tag{4}$$

where $g(\cdot)$ is a square-integrable function in the range to be chosen by the prioritization system, and random variable Y denotes the origin of the message. The definition of semantic distance is application-dependent and implicitly defined by users' preferences. For instance, users' metric for distance might be whether the author of a message is young, in which case $d(\boldsymbol{x}, \boldsymbol{z}, m) = \mathbb{1}\{\boldsymbol{z} \text{ is young}\}$. If the system aims to only distribute messages by young authors, then $g(d(\boldsymbol{x}, \boldsymbol{z}, m)) = \delta(d(\boldsymbol{x}, \boldsymbol{z}, m))$, where δ denotes the Dirac delta function; iff the author, is young, $\boldsymbol{x}$ sees the message with probability 1.

Unlike this model, Rangzen uses deterministic forwarding to maximize message propagation. Nonetheless, our model is useful for building intuition about the system while abstracting away complicated factors like storage limits and mobility patterns.

Definition The *spread* of the mapping g is defined as its second moment:

$$S(g(t)) = \int_0^\infty t^2 g(t) dt.$$

This definition captures the notion of how much 'unusual' content is served to nodes compared to regular, preferred content (where preferred content has a d-score closer to 0). Higher spread enables greater dissemination of content within the network, whereas lower spread enables more targeted, relevant content dissemination. One possible objective for a prioritization algorithm is to minimize spread.

$\boldsymbol{x}$'s a posteriori distribution of the identity of message author Y is described by $P(Y = y | \boldsymbol{x} = 1)$. Assuming the message is equally likely to originate from any node, this quantity is equivalent (up to a constant factor) to $P(\boldsymbol{x} = 1 | Y = y)$.

Definition The *anonymity* of the author in our model is defined as the Shannon entropy of Y given an observation $\boldsymbol{x} = 1$. Since $P(Y | \boldsymbol{x} = 1)$ is related to g by a constant factor, this entropy is equivalent to

$$A(\boldsymbol{z} | \boldsymbol{x} = 1) = h(g) = -\int_0^\infty g(s) \log g(s) ds.$$

where $h(\cdot)$ denotes the differential entropy function. Thus we maximize the author's anonymity by maximizing the entropy of g (normalized). Suppose we limit ourselves to prioritization functions that are step functions when normalized:

$$g(x) \propto \begin{cases} 1/t & : 0 \leq x \leq t \\ 0 & : \text{ Otherwise} \end{cases}$$

For this class of prioritization functions, it is easy to show that

$$h(g) = \log_2 \sqrt{3} + \frac{1}{2} \log_2 S(g). \tag{5}$$

(5) is monotonically increasing in $S(g)$, so the better the prioritization scheme (indicated by a lower spread $S(g)$), the worse the sender's anonymity (indicated by a lower entropy $h(g)$). This intuition, expressed in a more general form through the Cramér-Rao inequality (albeit for Fisher information rather than Shannon entropy), suggests that the trade-off between message prioritization and anonymity is fundamental to a broader class of broadcast communication systems; we feel this problem merits further study.

6. DISCUSSION

In this section we examine the our results and their implications, including the security properties of Rangzen's protocols, and discuss broader implications of our design.

6.1 Summary of Results

Our results indicate that Rangzen and Rangzen-like systems could continue to deliver predominantly legitimate messages during an Internet blackout while protecting the anonymity of message authors. Next we summarize the results in more detail.

Message propagation. In simulation, Rangzen delivered messages from honest nodes to over 80 percent of the population within 24-48 hours, depending on the prioritization noise parameters (Figure 1). Figure 3 indicates that messages from popular nodes may spread up to 33 percent more than those from adversarial nodes.

Robustness of the network. Figure 4 indicates that Rangzen is robust to localized denial of service attacks (e.g. jamming) when 10 percent of the population is an attacker, using devices with ranges up to 1.3 km. We believe such an attack to be beyond the capabilities of any adversary. However, Figure 3 suggests that coalitions of adversarial nodes cannot dominate network resources as long as they have few friends. A coalition of 6 adversarial nodes in a network of 400 nodes performed only marginally better on average than individual honest nodes selected uniformly at random.

Protection for users.

Authorship deniability. Rangzen allows users to deny authorship of any message with non-negligible probability (§5.3.1). §5.4 describes a trade-off we have identified between quality of prioritization (i.e. ability to weed out malicious messages) and anonymity.

Device Capture. If an adversary captures $\boldsymbol{b}$'s device, $\boldsymbol{b}$'s friend IDs are password-protected. Without input from $\boldsymbol{b}$, the adversary can only learn mutual friends using the chosen-input PSI-Ca attack. Even with $\boldsymbol{b}$'s password, friend IDs are not stored on device—only hashes of their IDs. The adversary could also conduct a message spreading attack with the stolen device, but $\boldsymbol{b}$'s friends, upon learning of his stolen device, would send revocation messages, gradually choking the stolen device from the network.

Device Impersonation. Signed revocation messages prevent spoofing. Impersonation only helps if the adversary has the client's friend list, which requires the true owner's password input.

Trust Graph Extraction. §5.3.2 shows that a resource-limited adversary cannot learn a significant portion of the trust graph. We amplify this effect by randomly adding and deleting inputs, and fixing the maximum inputs to the PSI-Ca protocol.

6.2 Protocol Security

Rangzen combines friend establishment and private set intersection of friend identities into a protocol for anonymous message dissemination. While we did not explore our prototype or systems challenges in this paper, we briefly discuss protocol security.

An attacker spreading propaganda must do so via the Rangzen protocol. A Rangzen user will only store a new message if the message is authored by the node itself or received during a peer encounter. Attackers that do not corrupt the client software must therefore attack the peer encounter protocol to spread messages.

Propaganda spread. Rangzen clients reject messages from peers with whom they cannot complete private set intersection. If the attacker performs the PSI protocol correctly, then their success at spreading propaganda relies on their ability to form trust relationships with users. Otherwise, the attacker can attempt to misbehave during the PSI protocol. Both points are discussed below.

Attacking friend management.

Friend addition. Rangzen clients do not add friend IDs to their database without an in-person exchange. Thus attackers must either capture devices or socially engineer targets to befriend honest users. If an attacker learns friend IDs, he can include them in his own friendship database, forming a directed edge in the graph. But since the protocol never shares friend identities, the attacker would have to befriend or harass a user to learn real identities.

Friend revocation. Since revocation messages are signed, an adversary cannot directly spoof them; at worst, the adversary can drop revocation messages to hinder their spread. Our simulations indicate that this is unlikely to be a problem; even in the presence of adversarial coalitions dropping honest messages, those messages were still able to reach most of the population.

Attacking trust computation.

Malicious PSI-Ca. The private set intersection with cardinality (PSI-Ca) protocol is only provably secure against semi-honest adversaries [20]. Adversaries could deploy a few tactics, none of which helps them gain trust or further their aims. Suppose an adversary node $\boldsymbol{a}$ meets an honest node $\boldsymbol{b}$. 1) If $\boldsymbol{a}$ refuses to participate in the protocol, $\boldsymbol{b}$ automatically discards $\boldsymbol{a}$'s messages, preventing the adversary from disseminating content. 2) If $\boldsymbol{a}$ executes the modular exponentiation incorrectly, $\boldsymbol{b}$ will find no mutual friends with high probability, thereby reducing $\boldsymbol{b}$'s trust for $\boldsymbol{a}$. 3) If $\boldsymbol{a}$ submits insecure friend ID encryptions (e.g. in a lower field), $\boldsymbol{b}$ will conduct the modular exponentiations in the proper field, hiding the information the adversary hoped to gain. 4) Nodes input a fixed number of entries to the PSI-Ca protocol to avoid the adversary learning information from the number of inputs. This also disincentivizes promiscuous trust, which reduces the damage from compromised devices.

Adversarial history. Adversaries can learn social graph edges by submitting only a single real ID friend to the PSI protocol per encounter. If the intersection is cardinality 1, the attacker learns that their communication partner is friends with that ID. We call this a chosen input attack on the trust computation. If an adversary can identify peers it meets (out-of-band or through fingerprinting), this would at most allow the adversary to learn as much of the graph as described in §5.3.2. By rate limiting encounters, honest nodes can restrict the amount of information leaked through this channel. Policies for this restriction are an area of future work on Rangzen.

Denial of Service. Attackers might attempt to launch denial of service attacks by overwhelming the system with messages, including revocation messages on the control plane. Neither storage nor bandwidth can be overwhelmed by such a flood of messages since the prioritization mechanism applies to all such messages.

7. CONCLUSIONS AND FUTURE WORK

While we find that Rangzen can enable legitimate anonymous messaging in the presence of active attacks, there are a number of avenues for future exploration. We do not deal with attackers who can effectively gain the trust of many users. Our friendship revocation protocol assumes that humans in the network are able to identify adversaries. If this assumption is false, Rangzen's performance is likely to suffer. However, we believe that solving this problem with technology is very difficult, and such an attack is not scalable even for powerful adversaries. However, further study is needed to understand what mobility or social patterns will break Rangzen.

Moving forward, we hope to investigate several of the issues we have identified, such as the trade-off between prioritization and anonymity, and Rangzen's robustness to jamming. We also will soon begin large-scale real-world tests of our Rangzen app in an adversarial-game environment. We hope to learn how effective the algorithm is in practice, but we also are keen to observe how people behave in adversarial environments. Finally, we would like to enable friend establishment without physical proximity, such as over the phone, using a combination of existing authentication approaches to defuse MITM risks.

8. REFERENCES

[1] Anonymizer. https://www.anonymizer.com/.

[2] Private Internet Access. https://www.privateinternetaccess.com/.

[3] Secret. https://secret.ly/.

[4] Tavern. https://tavern.com/.

[5] Twitter. https://www.twitter.com/.

[6] Syria crisis: Guide to armed and political opposition. *BBC News*, December 13, 2013.

[7] Libya protests: 84 killed in growing unrest, says HRW. *BBC News*, February 19, 2011.

[8] Why is Ukraine in turmoil? *BBC News*, February 22, 2014.

[9] Greece protest against austerity package turns violent. *BBC News*, June 28, 2011.

[10] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.

[11] G. Bigwood, D. Rehunathan, M. Bateman, and S. Bhatti. CRAWDAD data set st_andrews/sassy (v. 2011-06-03). Downloaded from http://crawdad.org/st_andrews/sassy/, June 2011.

[12] C. Buckley and R. Donadio. Buoyed by Wall St. Protests, Rallies Sweep the Globe. *New York Times*, October 16, 2011.

[13] R. Canetti, C. Dwork, M. Naor, and R. Ostrovsky. Deniable encryption. In *Proceedings of CRYPTO*, 1997.

[14] D. Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of cryptology*, 1(1), 1988.

[15] E. Cho, S. Myers, and J. Leskovec. Friendship and Mobility: Friendship and Mobility: User Movement in Location-Based Social Networks. In *ACM KDD*, 2011.

[16] M. Chulov. Syria shuts off internet access across the country. *The Guardian*, November 29, 2012.

[17] A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4), 2009.

[18] H. Corrigan-Gibbs and B. Ford. Dissent: accountable anonymous group messaging. In *Proceedings of ACM CCS*, 2010.

[19] J. Cowie. Egypt Leaves the Internet. *Renesys*, January 2011. http://www.renesys.com/2011/01/egypt-leaves-the-internet/.

[20] E. De Cristofaro and G. Tsudik. Practical private set intersection protocols with linear complexity. In *Proceedings of Financial Cryptography*, 2010.

[21] M. Dehmer and A. Mowshowitz. A history of graph entropy measures. *Information Sciences*, 181(1), 2011.

[22] C. Diaz, C. Troncoso, and A. Serjantov. On the impact of social network profiling on anonymity. In *Proceedings of PETS*, 2008.

[23] J. Diaz. Iran Shuts Down Google, Will Completely Cut Citizens Off the Internet. *Gizmodo*, September 24, 2012.

[24] R. Dingledine, M. J. Freedman, and D. Molnar. The free haven project: Distributed anonymous storage service. In *Designing Privacy Enhancing Technologies*, 2001.

[25] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. In *Proceedings of USENIX Security*, 2004.

[26] C. Dong, L. Chen, and Z. Wen. When private set intersection meets big data: An efficient and scalable protocol. In *Proceedings of ACM CCS*, 2013.

[27] J. R. Douceur. The sybil attack. In *Proceedings of IPTPS*, 2002.

[28] K. Fahim. Violent Clashes Mark Protests Against Mubarak's Rule. *New York Times*, January 26, 2011.

[29] A. Fantz. Son: Iranian dad arrested for my facebook posts. *CNN*, July 12, 2012.

[30] F. Fassihi. Iranian Crackdown Goes Global. *Wall Street Journal*, December 3, 2009.

[31] B. Ford, J. Strauss, C. Lesniewski-Laas, S. Rhea, F. Kaashoek, and R. Morris. Persistent personal names for globally connected mobile devices. In *Proceedings of USENIX/ACM OSDI*, 2006.

[32] M. J. Freedman and R. Morris. Tarzan: A peer-to-peer anonymizing network layer. In *Proceedings of ACM CCS*, 2002.

[33] C. Garman, M. Green, and I. Miers. Decentralized anonymous credentials. In *Proceedings of ISOC NDSS*, 2014.

[34] J. Geddes, M. Schuchard, and N. Hopper. Cover your ACKs: pitfalls of covert channel censorship circumvention. In *Proceedings of ACM CCS*, 2013.

[35] M. Grossglauser and D. Tse. Mobility increases the capacity of ad-hoc wireless networks. In *IEEE INFOCOM*, 2001.

[36] S. Hasan, Y. Ben-David, G. Fanti, E. Brewer, and S. Shenker. Building dissent networks: Towards effective countermeasures against large-scale communications blackouts. In *Proceedings of the USENIX FOCI Workshop*, 2013.

[37] C. Hedges. *Death of the liberal class*. Nation Books, 2010.

[38] T. Isdal, M. Piatek, A. Krishnamurthy, and T. Anderson. Privacy-preserving P2P data sharing with OneSwarm. In *Proceedings of ACM SIGCOMM*, 2010.

[39] R. Jansen and R. Beverly. Toward anonymity in delay tolerant networks: threshold pivot scheme. In *IEEE MILCOM*, 2010.

[40] G. Y. Keung, Q. Zhang, and B. Li. The base station placement for delay-constrained information coverage in mobile wireless networks. In *Proceedings of IEEE ICC*, 2010.

[41] J. O. Koehler. *STASI: The untold story of the East German secret police*. Basic Books, 1999.

[42] G. Kossinets and D. J. Watts. Empirical analysis of an evolving social network. *Science*, 311(5757):88–90, 2006.

[43] M. Li, N. Cao, S. Yu, and W. Lou. Findu: Privacy-preserving personal profile matching in mobile social networks. In *Proceedings of IEEE INFOCOM*, 2011.

[44] X. Liang, X. Li, K. Zhang, R. Lu, X. Lin, and X. Shen. Fully anonymous profile matching in mobile social networks. *IEEE JSAC*, 2013.

[45] X. Lu, P. Hui, D. Towsley, J. Pu, and Z. Xiong. Anti-localization anonymous routing for delay tolerant network. *Computer Networks*, 54(11), 2010.

[46] S. Luke, C. Cioffi-Revilla, L. Panait, K. Sullivan, and G. Balan. Mason: A multiagent simulation environment. *Simulation*, 81(7), 2005.

[47] J. McAuley and J. Leskovec. Learning to discover social circles in ego networks. In *Proceedings of NIPS*, 2012.

[48] P. Meroni, S. Gaito, E. Pagani, and G. P. Rossi. CRAWDAD data set unimi/pmtr (v. 2008-12-01). Downloaded from http://crawdad.org/unimi/pmtr/, Dec. 2008.

[49] I. Miers, C. Garman, M. Green, and A. D. Rubin. Zerocoin: Anonymous distributed e-cash from bitcoin. In *Proceedings of IEEE Security and Privacy*, 2013.

[50] M. Nagy, E. De Cristofaro, A. Dmitrienko, N. Asokan, and A.-R. Sadeghi. Do I know you?: efficient and privacy-preserving common friend-finder protocols and applications. In *Proceedings of ACM Computer Security Applications Conference*, 2013.

[51] A. Narayanan and V. Shmatikov. De-anonymizing social networks. In *Proceedings of IEEE Security and Privacy*, 2009.

[52] R. Needleman. Firechat network-free chat could be big. and now it's on android. *Yahoo News*, April 3, 2014.

[53] G. Noubir. On connectivity in ad hoc networks under jamming using directional antennas and mobility. In *Wired/Wireless Internet Communications*. 2004.

[54] A.-K. Pietilainen. CRAWDAD data set thlab/sigcomm2009 (v. 2012-07-15). Downloaded from http://crawdad.org/thlab/sigcomm2009/, July 2012.

[55] M. Piorkowski, N. Sarafijanovic-Djukic, and M. Grossglauser. CRAWDAD data set epfl/mobility (v. 2009-02-24). Downloaded from http://crawdad.org/epfl/mobility/, Feb. 2009.

[56] A. Post, V. Shah, and A. Mislove. Bazaar: Strengthening user reputations in online marketplaces. In *Proceedings of USENIX/ACM NSDI*, 2011.

[57] N. Rashevsky. Life, information theory, and topology. *The bulletin of mathematical biophysics*, 17(3), 1955.

[58] F. Reid and M. Harrigan. An analysis of anonymity in the bitcoin system. In *Security and Privacy in Social Networks*. 2013.

[59] M. K. Reiter and A. D. Rubin. Crowds: Anonymity for web transactions. *ACM TISSEC*, 1(1), 1998.

[60] C. Rhoads and G. Fowler. Egypt Shuts Down Internet, Cellphone Services. *The Wall Street Journal*, January 29, 2011.

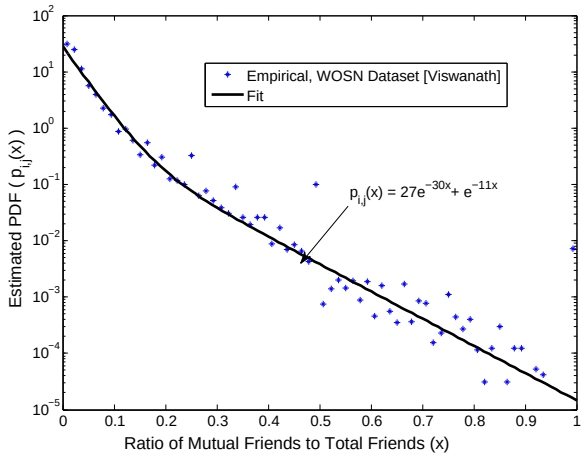

Figure 9: Estimated probability density function of the ratio $p_{i,j}$ over all node pairs. A sum of exponentials over the range $x \in [0, 1]$ models the data in [64] well. This model of pairwise trust is used to estimate anonymity set sizes.

[61] B. Schneier. The Battle for Power on the Internet. *The Atlantic*, October 24, 2013.

[62] N. Shachtman. Syria's internet blackout explained. *Wired*, November 30, 2012.

[63] Z. Tufekci. After the Protests. *New York Times*, March 20, 2014.

[64] B. Viswanath, A. Mislove, M. Cha, and K. P. Gummadi. On the evolution of user interaction in facebook. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Social Networks*, August 2009.

[65] B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove. An analysis of social network-based sybil defenses. In *Proceedings of ACM SIGCOMM*, 2011.

[66] S. Walker and O. Grytsenko. Text messages warn Ukraine protesters they are 'participants in mass riot'. *The Guardian*, January 21, 2014.

[67] D. I. Wolinsky, E. Syta, and B. Ford. Hang with your buddies to resist intersection attacks. In *Proceedings of ACM CCS*, 2013.

[68] R. Worth and N. Fathi. Violent Clashes Mark Protests Against Mubarak's Rule. *New York Times*, June 14, 2009.

[69] H. Yu, M. Kaminsky, P. B. Gibbons, and A. Flaxman. Sybilguard: defending against sybil attacks via social networks. In *Proceedings of ACM SIGCOMM*, 2006.

[70] X. Zhang, G. Neglia, J. Kurose, and D. Towsley. Performance modeling of epidemic routing. *Computer Networks*, 51(10), 2007.

APPENDIX

A. REVOCATION AND RENEWAL

While not central to this paper, here we briefly describe how we allow users to revoke trust credentials if they no longer trust a particular node (e.g., due to shifting coalitions or a friend's device being compromised). If Alice no longer trusts Bob, she can manually remove Bob's ID by entering her own password and Bob's name. However, Bob still has Alice's ID, which he can exploit to spread his own messages. Alice therefore creates a new ID for herself. She encrypts the new and old IDs with multicast encryption using her friends' IDs as keys (except for Bob), and propagates this control plane message. When a friend of Alice's receives the message, she can decrypt and update Alice's ID. Control plane messages obey the same rules as regular content—our prioritization algorithm is expected to prevent an adversary from jamming such messages or flooding the system with spurious messages.

B. ANONYMITY SET MODELING

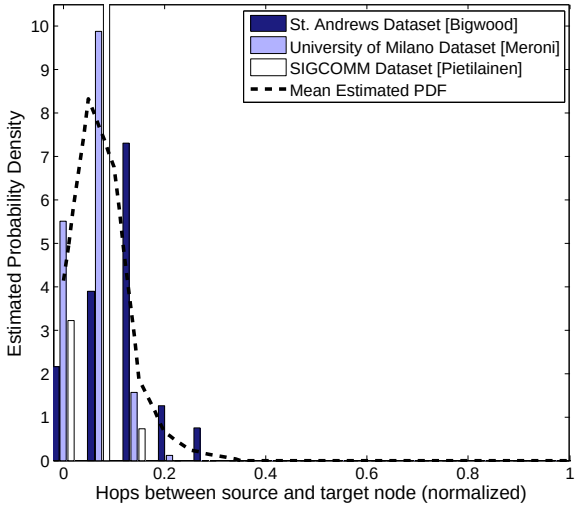

Figure 10: Empirically-estimated PDF for the minimum number of hops between pairs of nodes in mobility traces [11,48,54]. Most nodes are within a few hops of one another. We use the mean CDF measurement to model $P(N \leq n|\Omega)$.

A message's priority score S depends on the number of hops the message took. In particular, we can define the received priority after $N = n$ hops (S_n) recursively as follows:

$$\begin{aligned} S_n &= p_n \cdot S_{n-1} + z_n \\ S_0 &= 1 \end{aligned}$$

where z_i is the noise added by the ith node, and p_i is the priority score at the ith node. z_i's distribution is designed, but priority scores p_i depend on graph and mobility properties.

This priority depends on $p_{i,j}$, the scaling factor when messages pass from j to i, as defined in Equation 1. Empirical evidence shows that degree distributions in social networks obey a power law [17]. We found that in the Facebook WOSN dataset [64], mutual degree distributions also obey a power law tail distribution, but the *ratio* p_{ij} across *all node pairs* appears to be better-modeled by a truncated sum of exponentials (Figure 9).[10] This trust metric is not heavy-tailed, so the fraction of nodes with highly overlapping friend sets is vanishingly small. This motivates the sigmoid in equation 1, which assigns high trust even if nodes don't share a large fraction of mutual friends.

We estimated $P(N = n|\Omega)$ empirically from several datasets (Figure 10). For every pair of nodes in the dataset (i, j), we measured the minimum number of times a message would need to be forwarded before reaching target j from source i. This measurement enables a lower bound estimate on how many hops in the (time-varying) connectivity graph separate an arbitrary message from its creator. Figure 10 illustrates these measurements, normalized by the total network nodes. Most pairs are a few hops apart, in part due to the small scale of these mobility datasets.

[10]Technically, this probability is only defined over rational values, but we approximate the function as having a continuous domain.