

GSFord: Towards a Reliable Geo-Social Notification System

Kyungbaek Kim

Dept. of Electronics and Computer Engineering,
Chonnam National University, South Korea
Email: kyungbaekkim@jnu.ac.kr

Ye Zhao and Nalini Venkatasubramanian

Dept. of Computer Science
University of California, Irvine, USA
Email: {yez,nalini}@ics.uci.edu

Abstract—The eventual goal of any notification system is to deliver appropriate messages to all relevant recipients with very high reliability in a timely manner. In particular, we focus on notification in extreme situations (e.g. disasters) where geographically correlated failures hinder the ability to reach recipients inside the corresponding failed region. In this paper, we present GSFord, a reliable geo-social notification system that is aware of (a) the geographies in which the message needs to be disseminated and (b) the social network characteristics of the intended recipient, in order to maximize/increase the coverage and reliability. GSFord builds robust geo-aware P2P overlays to provide efficient location-based message delivery and reliable storage of geo-social information of recipients. When an event occurs, GSFord is able to efficiently deliver the message to recipients who are either (a) located in the event area or (b) socially correlated to the event (e.g. relatives/friends of those who are impacted by an event). Furthermore, GSFord leverages the geo-social information to trigger a social diffusion process, which operates through out-of-band channels such as phone calls and human contacts, in order to reach recipients which are isolated in the failed region. Through extensive evaluations, we show that GSFord is reliable; the social diffusion process enhanced by GSFord reaches up to 99.9% of desired recipients even under massive geographically correlated regional failures. We also show that GSFord is efficient even under skewed distribution of user populations.

Keywords—Event Notification, Geographical Failure, Social Diffusion

I. INTRODUCTION

The eventual goal of an event notification system is to deliver appropriate messages to all relevant recipients with very high reliability in a timely manner. We are motivated by the case of disaster alerting and warning systems where a notification system alerts impacted populations on how to take self-protective action and prevent loss of lives and property. In extreme situations, the eventual coverage/reliability of the message dissemination is of utmost importance.

Firstly, in natural or human-induced disasters, information needs are strongly correlated to the geographical location of the event. For instance, severe ground shaking and damage occur in areas that are within close geographic proximity to an earthquake epicenter; Tornadoes are likely to destroy infrastructures in areas that lie along their damage paths. Ideally, a notification system should exploit this geographical correlation to warn populations in the disaster region as soon as possible.

Secondly, there exists social correlation of information needs. Individuals are often interested in receiving accurate information on disasters (occurrence and progress) which may affect their loved ones and desire detailed information on the current status of their relatives and friends. They are interested in receiving the information even if they are not physically located in the disaster region. For example, if a forest fire occurs near an elementary school, parents will need to be informed of the current situation as soon as possible; information on the current status of their children

will help them make decisions on how to respond quickly and effectively (e.g. plan details on how families will evacuate).

Geo-correlated failure propagation caused by disasters adds to the complexity of event notification. For example, when an earthquake occurs, one can expect catastrophic network failures or blackouts in the affected region[1]. Recipients inside and outside the region can be affected by the failure of the dissemination infrastructure. Scalability also poses additional challenges - as the number of participants increase, personalized warning can consume significant resources (e.g. bandwidth, processing power) and possible bottlenecks can result in delays and warning inconsistencies.

Typical approaches to disaster alerting include sirens, subscription-based warnings from a centralized broker (USGS Shakecast, SMS Gateways) and cell broadcasts (currently implemented in parts of Asia and Europe). These systems lack information specificity by providing a broadcasting service to the entire set of participating recipients. Application layer multicasting (ALM)[6], [5], [9] can be used to multicast specific messages to a fine-grained set of recipients, but these systems are not aware of the geography. Subscription based systems[12] and geography aware P2P systems[13], [14], [15] have been proposed for dynamic geography aware data retrieval or dissemination. However, these efforts did not consider the need to reach socially correlated remote recipients who are currently not located in the disaster region. Additionally these efforts lack explicit fault tolerance against the geographically correlated regional failures[7], [8].

In this paper, we present GSFord, a geo-social notification system that is aware of (a) the geographies in which the message needs to be disseminated and (b) the social network characteristics of the intended recipients. GSFord builds self-organizing, reliable, and fault-tolerant structured overlays, which adjust the structure dynamically based on the population of the geography and replicate the information of the overlay to tolerate geographically correlated regional failures. Section III describes the techniques for constructing GSFord overlays that support efficient event notification under the non-uniform user distribution; the overlays also provide reliable storage of social network information under extreme regional failures. Section IV describes the event dissemination processes in GSFord; we show that GSFord is able to efficiently disseminate event messages to dynamically defined recipients who are either geographically or socially correlated to the event. Under geographically correlated regional failures, GSFord reliably delivers messages to unfailed recipients outside the failed region through its online dissemination structure. To better reach those in failed regions, GSFord exploits a targeted social diffusion process triggered by using the stored social information. The social diffusion process enables propagation of event messages towards expected recipients through diverse out-of-band communication channels.

We evaluate GSFord through extensive simulations (Section

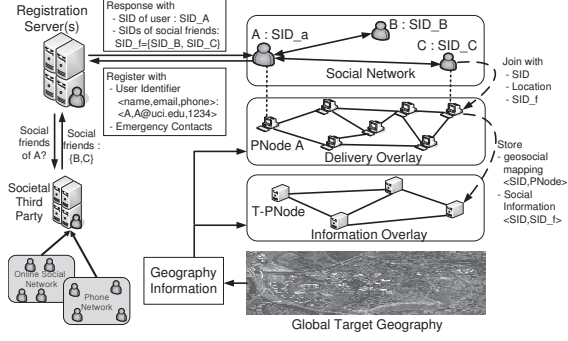


Fig. 1. The overview of GSFord.

V) in terms of reliability and efficiency. We show that the importance of preserving the social information of recipients in GSFord to maximize the eventual coverage of the event notification system, especially in the presence of large geographically correlated failures. Specifically, once GSFord ensures that socially correlated recipients receive the initial event message, the coverage of GSFord increases by around 15%; additionally around 90% of nodes experience a decrease in dissemination latency of up to 80%. Our results also demonstrate that the GSFord overlay provides efficient and reliable geo-aware regional multicasting with reasonable network overhead. In particular, the GSFord overlay maintains its performance even though the large-scale users are non-uniformly distributed.

II. GSFORD APPROACH

Figure 1 provides an overview of GSFord. GSFord operates over a **global target geography** (GTG) that defines the global geographical region over which geo-social notifications occur. We begin by describing some basic concepts, notations and assumptions used in GSFord to capture the social and geographical aspects for efficient and effective notification.

A. Social concepts of GSFord

The eventual recipients of notifications in GSFord are real users located inside the global target geography interested in receiving the notification and socially connected to each other using disparate communication channels. More formally, the set of users in the GTG form a social network graph, $G := (V, E)$, where a user $u \in V$ is connected to another user $v \in V$ if there exists a social link $e_{u,v} \in E$. In the above case, v is referred to as a **social acquaintance** of u . Let $A(u)$ denote the set of social acquaintances of u , so

$$A(u) = \{v : v \in V, e_{u,v} \in E\}$$

The **degree centrality** of a user is defined as the number of social acquaintances he/she has. We use $C_D(u)$ to denote the degree centrality of a user u , so

$$C_D(u) = |A(u)|$$

There is a subset of users in $A(u)$ who are closely tied to u in a society, such as family members, and therefore highly interested in the current status of u when disasters and the ensuing event notification process occurs. They are referred to as the **social friends** of u . We denote the set of social friends of u as $F(u)$ where $F(u) \subseteq A(u)$.

$F(u)$ and $A(u)$ can be obtained in many ways, such as direct user input, e.g. the list of emergency contacts that organizations usually require individuals to provide. It may also be obtained by profiling and analyzing a known social network

graph, such as online social network graphs from Facebook, Twitter or even phone network graphs from elementary school offices, based on the connectivity of social entities and the frequency of correspondence [18].

In GSFord, every user has a unique **public Social ID (SID)** that hides his real identity. We denote a user u 's SID as SID_u . We assume the presence of a **registration server**, which resides in an authorized domain, can generate unique SIDs for prospective users based on their personal information (name and contacts) and manage the mapping between a user real identity and his/her SID. During the registration, the server also transforms the list of social friends of a user u into a list of SIDs, say $\{SID_v : v \in F(u)\}$, by referring to their mappings. We also assume that the registration server is not a bottleneck since (a) registration and mapping is a one-time process, and (b) it is usually performed apriori, i.e. before notification; issues of surge and overload at the registration server are out of scope of the paper.

B. Geographical concepts of GSFord

GSFord has prior knowledge of the GTG; For simplicity, this is mapped into a 2 dimensional rectangular region, denoted as $GTG := ((0, 0), (x_{max}, y_{max}))$, where $(0, 0)$ and (x_{max}, y_{max}) are the coordinates for the bottom-left and top-right corners of the rectangular region respectively. A user executes the GSFord application on a personal device; we call the logical host on which the user resides as a **GSFord Physical Node (PNode)**, and denote the PNode for a user u as PN_u . We assume that a PNode has knowledge of its location (using positioning technologies such as GPS, WiFi fingerprinting, etc.), and the location is mapped to a point in the GTG, denoted as $Loc(PN_u) := (x, y)$. The PNode, PN_u , is used by the user to join and receive notifications from GSFord. Moreover, it maintains the user's geo-social information, which includes its current location $Loc(PN_u)$, the user's public social ID SID_u , the user's degree centrality $C_D(u)$ and the list of SIDs of his social friends $\{SID_v : v \in F(u)\}$.

We select a subset of PNodes that are more trusted and reliable and refer to them as **Trusted PNodes (T-PNodes)**. T-PNodes typically correspond to users that represent public figures and authorities, e.g. desktop machines governed by organizations such as local government agencies (fire, law enforcement) and university/company administrative authorities. T-PNodes are few (compared to the total number of users) and are maintained by authorized entities at their local sites. Furthermore, T-PNodes are uniformly distributed across the global target geography.

The message of an event is tied to (intended for) a specific sub region inside the GTG. For example, the alerting message of a campus fire is tied to the campus and its proximity. This region is called the **Possibly Affected Region (PAR)** of the message - for simplicity, we assume this region is rectangular as well and is denoted as PAR where $PAR \subseteq GTG$. Crisis events (e.g. earthquakes, tornadoes) may damage the communication infrastructures inside the PAR and cause a geographically correlated regional failure. We define the regional geographical failure as the **Possibly Damaged Region (PDR)**, denoted as PDR ; we assume $PDR \subseteq PAR$.

We consider the set of users inside a PAR as the **geographically correlated target recipients** of an event message, and the set is formulated as:

$$TR_g = \{u : u \in V, Loc(PN_u) \text{ in } PAR\}$$

As we mentioned earlier, the social friends of the geographically correlated target recipients are also possible recipients

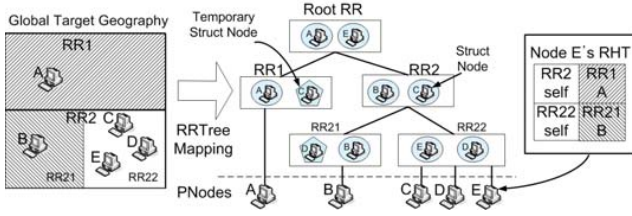


Fig. 2. The mapping between the global target geography and the RRTree-based overlay structure. $Th_s = 4$, $Th_{str} = 2$, and $N_c = 2$. $RR1$, $RR21$ and $RR22$ are leaf RRs

of the message. This set of users are called as the **socially correlated target recipients** and formulated as:

$$TR_s = \{v : v \in F(u), \forall u \in TR_g\}$$

Note that, unlike TR_g , users in TR_s may be located outside the PAR.

C. Overlays of GSFord

Using the concepts delineated above, we develop two distinct overlays - **Delivery Overlay** and **Information Overlay** in GSFord. The delivery overlay aims to reach PNodes associated with a given region (e.g. a PAR) efficiently and effectively (section III-B). On the other hand, the main purpose of the information overlay is to capture and maintain the geo-social information of participating users despite extreme damage (section III-C). To enable this, the information overlay replicates the stored information by using the conjugate region-based replication technique (section III-C1). Moreover, to mitigate security/privacy concerns, such as revealing personalizing information (e.g. social friends), the information overlay is constructed/maintained at only T-PNodes.

D. Event notification over GSFord

The delivery and information overlays are created apriori given knowledge of the geography and social connections of individual users. When an event occurs, a new message related to PAR is created. The message is conveyed to the geographically correlated target recipients (TR_g) by the delivery overlay. To reach the socially correlated target recipients (TR_s), the information overlay explores the stored geo-social information to identify and locate them and forwards the message via the delivery overlay.

One of the key objectives of GSFord event notification is to reach users in the extremely damaged region affected by a disaster who are less likely to be reachable via the delivery overlay. For this purpose, GSFord encourages the socially correlated target recipients (TR_s) to initiate targeted social diffusion to propagate the received message further by customizing the contents of the forwarded message (section IV).

III. GSFORD OVERLAYS : RRTREE-BASED GEO-SOCIAL AWARE OVERLAYS

A. RRTree-based geo-aware P2P overlay structure

To generate and maintain the delivery and information overlay, we develop a peer-oriented geo-aware multicast overlay structure called Responsible Region Tree (RRTree) such that (a) it is self-organizing and fault-tolerant, (b) it supports efficient geographical regional multicasting, even with non-uniformly distributed PNodes, and (c) it supports DHT-style reliable storage of social information under extreme geographically correlated regional failures.

The RRTree overlay structure is a hierarchically nested logical structure built by using the concept of a Responsible

Region (RR) which is a rectangular region inside the GTG (see Figure 2). The RRTree structure inherently maintains the following property. A parent RR subsumes the regions covered by its child RRs . Furthermore, the region covered by the root RR (i.e. GTG) can be obtained by a direct union of the regions covered by the leaf RRs . The RRTree structure adapts to the dynamics of user population and distribution through the process of RRTree growth and RRTree shrinkage, which handles the non-uniformly distributed PNodes gracefully.

Since the RRTree is a logical structure, information about an RR is basically stored at multiple PNodes located inside the RR ; we call these PNodes **Struct Nodes** to indicate that they are structure maintenance PNodes. In order to sustain both random overlay failures and geographically correlated regional overlay failures, there are at least Th_{str} struct nodes of a RR and they are sparsely distributed over the RR .

To achieve efficient routing in a RRTree overlay structure, we introduce the concept of a **Region Hopping Table** (RHT), maintained at each PNode that allows messages to hop to other non-overlapping regions without ascending/descending the RRTree. Unlike the RRTree which only has knowledge of its immediate neighbors (parent and child RRs), the RHT of a PNode covers the GTG . If PNodes are uniformly distributed, a RHT of a PNode has an average of $\log_{N_c} N$ rows, where N is the total number of PNodes. According to this, a point-to-point routing with the RHT takes an average of $\log_{N_c} N$ hops.

The details of the RRTree-based geo-aware P2P overlay structure are described in [20].

B. Delivery overlay: Efficient regional multicasting

The delivery overlay (DOv) is composed of PNodes and constructed by using the RRTree-based overlay construction method. The purpose of the delivery overlay is to support efficient and effective geographical regional multicasting to all of the PNodes inside a given region. To start a regional multicasting of a message, a PNode initiates a message in the format of $M = [T, CT, MP]$, where T represents eventual target region of the message, $CT (\in T, \text{and initially } CT = T)$ indicates the target region for the next immediate forwarding, i.e. where the current PNode wishes to forward the message and MP is the message payload. Note that the aim of the regional multicasting is that the message M must be eventually forwarded to all of the PNodes corresponding to T .

In the delivery overlay, a PNode mainly uses the regional multicasting with RHT to forward a message, while the multicasting with RRTree is only used for the fail over purpose. That is, a PNode finds out the subsequent PNodes which need to convey the message by selecting the routing table entries in its RHT whose corresponding RRs partially overlaps with CT . Before forwarding the message to a next PNode of a RR , CT is updated to cover the region of overlap between the original CT in the message and the RR ; this is done to prevent in unnecessary message forwarding. The multicasting continues until a leaf RR is encountered which subsumes CT .

However, while RHT supports efficient geographical regional multicasting, information in the RHT may become stale under overlay failures. As a result, we may not reach contact PNodes in a staled entry; furthermore the stale information can cause loops in the routing. The unreachable PNodes are easily detected when trying to forward a message. To detect a routing loop, the message needs to piggyback the previous routing path. Whenever the stale RHT information is detected, the message is forwarded through RRTree, and the RHT replaces the stale contact PNodes with the newly obtained information from RRTree.

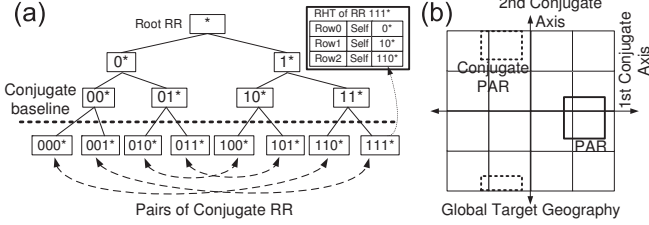


Fig. 3. (a) Example of region ID assignment and conjugate RR.; (b) Example of conjugate PAR with two conjugate axes.

C. Information overlay: Geo-Social information storage

The purpose of the information overlay is to reliably store geo-social information of GSFord users and to identify and locate the set of the socially correlated target recipients (TR_s) of a message during event notification (see section IV). The information overlay is managed using the RRTree-based overlay construction method too. RRs in the information overlay correspond to geographic regions, and the root RR of the information overlay is the same as the root RR of the delivery overlay which covers the GTG . A leaf RR in the information overlay is associated with the T-PNodes whose locations are inside the leaf RR . Note that the information overlay is composed of only T-PNodes rather than over all PNodes in order to alleviate the privacy concern.

In order to store geo-social information of a PNode, the PNode, PN_u , finds out the corresponding leaf RR in the information overlay, that subsumes its location. Then, PN_u provides its geo-social information, in the form of $[PN_u, Loc(PN_u), SID_u, C_D(u), \{SID_v : v \in F(u)\}]$ to the T-PNodes associated with the leaf RR . Eventually, the T-PNodes (i.e. IOv) store the geo-social information of all current online PNodes in GSFord.

To retrieve a user's (PNode's) location from his SID, the information overlay implements a geo-social mapping function. The geo-social mapping is a DHT-style mapping of $SIDs$ to region IDs ($RIDs$). A region ID is a unique identifier (bit vector) for a leaf RR in the information overlay and has a fixed number of bits (160 in our case). A RID assigned dynamically during splitting/merging of a RR as illustrated in Figure 3(a) to guarantee that the ID ranges covered by leaf RR are non-overlapping and these ranges together cover the entire ID space. The RID of the root RR is *, a *wildcard mask symbol*, which means it matches any values in the whole ID space.

A SID is hashed (using the SHA-1 hash function) to yield a bit vector, $Hash(SID)$, of the same length as RID . The geo-social mapping of a SID is stored in the T-PNode corresponding to the leaf RR whose RID has the best prefix-match (using bitwise comparison) to $Hash(SID)$. The SHA-1 hash function generates uniformly distributed hashed value w.h.p., and the storage/query load of the geo-social mapping over the entire information overlay can be balanced. To find the leaf RR corresponding to SID_u , we can conduct the RHT/RRTree routing process with slight modified routing tables as illustrated in Figure 3(a).

1) *Conjugate region based replication*: To cope with the information losses and inaccessible networks due to extreme geographically correlated regional failures, the information overlay replicates the stored information by using the *conjugate region based replication*. The **conjugate region** (R_C) of a region (R) is a geographically distant region to R and less likely impacted by the regional failure related to R . The conjugate region concept is inspired by the popular "out-of-state relative" concept used for emergency contacts. To do

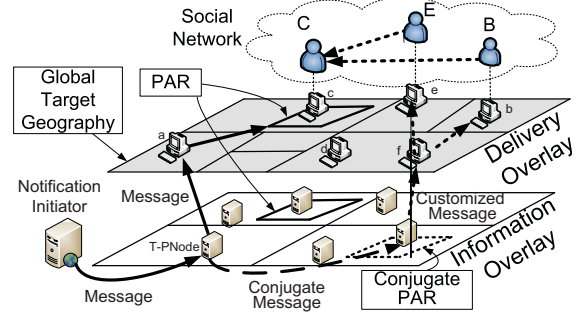


Fig. 4. An example of the notification process of GSFord.

this, the mapping between R and R_C ($f_C : R \rightarrow R_C$) should satisfy the following properties: (a) the mapping is known apriori by GSFord and easily maintained, (b) R_C should be less likely to fail simultaneously with R under a geographical regional failure, and (c) the geographical distance of any R , R_C pair should be similar in order to ensure fairness of information reliability.

To satisfy the desired properties of conjugate region, we come up with a simple mapping function as

$$\begin{aligned} f_C(x, y) &= (x_C, y_C) \\ &= ((x + x_{max}/2) \% x_{max}, (y + y_{max}/2) \% y_{max}) \end{aligned}$$

By using f_C , a region $((x_1, y_1)(x_2, y_2))$ can find its conjugate region $((x_{1C}, y_{1C})(x_{2C}, y_{2C}))$, which might be a wraparound region in GTG as illustrated in Figure 3(b). According to f_C , we can virtually draw two conjugate axes, and these two axes divide GTG into four quadrant virtual regions. The top-left region and the top-right region are matched to the bottom-right region and the bottom-left region as the conjugate region, respectively. The concept of the conjugate region can also be interpreted using the region ID concept as illustrated in Figure 3(a). We can use two bits to divide GTG into four quadrant virtual regions. The conjugate region ID of a region ID is obtained by conducting XOR bit operation of first 2 bits of the region ID.

IV. EVENT NOTIFICATION

Key intuitions behind efficient event notification in GSFord include (a) reaching recipients in geographically correlated target recipients, TR_g , who can propagate the message further via the delivery overlay and (b) leveraging reachable recipients in socially correlated target recipients, TR_s , to forward the message to recipients in the PDR through social diffusion over **diverse out-of-band communication channels**. Moreover, the observations about the social diffusion process suggest customizing the social diffusion process by (a) selecting good initiators and (b) modifying content of the messages, can help trigger more accurate and enable faster targeted social diffusion.

The overall of notification process of GSFord is illustrated in Fig. 4. The Three distinct steps include (a) reaching TR_g using the delivery overlay; (b) obtaining TR_s from the information overlay; and (c) message content customization and delivery of the customized messages to PNodes in TR_s .

Reach geographically correlated target recipients: Let us assume that an event message $M = [PAR, CT, MP]$ is generated by an authorized source (e.g. USGS, local government) communicated to a T-PNode in GSFord. This initial T-PNode starts the geographical regional multicasting of M to reach TR_g by using the delivery overlay.

Obtain socially correlated target recipients: At the same time, in order to reach TR_s , the initial T-PNode generates a conjugate message of M , $M_C = [PAR_C, CT, MP]$, where PAR_C is the conjugate region of PAR , and forwards M_C to T-PNodes corresponding to PAR_C via the information overlay. Upon receiving M_C , each of the T-PNodes in PAR_C performs a lookup to determine socially correlated recipients (TR_s) of the message in the form of $SIDs$. Their corresponding PNodes of TR_s are obtained by querying the geo-social mapping of the obtained $SIDs$.

Customize message content: After determining TR_s and their locations, T-PNodes may modify the original content of the message asking users in TR_s to contact social friends in the PAR/PDR . We develop 3 levels of customization with increasing degree of complexity: *coarse*, *PAR-targeted*, and *PDR-targeted*. As the complexity of the customization increases, triggering the more targeted and more aggressive social diffusion is expected.

COARSE customization : The intuition here is that the high centrality users can potentially diffuse the received message to a larger number of users in the PAR . To do this, T-PNodes pick top- K users in PAR with highest degree centrality. The customized message MP_u for a user u is $MP_u = MP + \{SID_v\}$ where v are u 's social friends among the top- K users in PAR , i.e. $v \in F(u) \cap TK_{CD}(PAR)$.

PAR-TARGETED customization : The PAR-targeted customization indicates the social friends to be contacted more specifically. That is, the customized message is $MP_u = MP + \{SID_v\}$ where v are u 's social friends located in PAR , i.e. $v \in F(u), PN_v \in PAR$.

PDR-TARGETED customization : In the case that PDR of an event is correctly defined, T-PNodes can customize the message much more precisely. The customized message is $MP_u = MP + \{SID_v\} + FLAG_{PDR}$ where v are u 's social friends located in PDR , i.e. $v \in F(u), PN_v \in PDR$. $FLAG_{PDR}$ indicates that MP_u needs to be handled as an *urgent* message in order to encourage users to contact the social friends in PDR by all means as soon as possible.

After customization, the individually modified message, MP_u , is forwarded to the PNode of a user u through the delivery overlay. Once the GSFord client on the PNode receives the message, it translates the list of $SIDs$ into human-readable form (e.g. names of friends) based on its local information. If the customized message has the flag of $FLAG_{PDR}$, the GSFord client phrases the message content as *urgent*.

V. EVALUATION

A. Evaluation settings

To simulate both of the propagation of messages over the geography-aware delivery overlay of GSFord and the social diffusion of messages over a social network graph, we mapped a sampled social network graph (50K nodes and 880K edges) [19] into the global target geography, GTG , which is set as a 131K by 131K meters square region. Each node in the social network graph plays as each PNode. PNodes are distributed by using uniformly random distribution and non-uniform distribution. For the case of non-uniform distribution, we use a truncated Gaussian distribution with $\mu = 45K$ meters and $\sigma^2 = 45K$ meters, and this setting mimics the demographics of the Southern California region including Los Angeles County, Orange County and Riverside County[2]. For the information overlay, we use 100 T-PNodes distributed uniformly at random.

The parameters for message dissemination in the delivery overlay such as latency and context switch overhead obtained

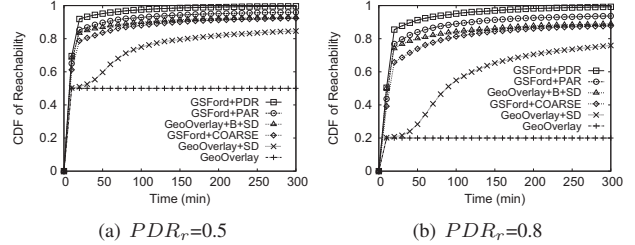


Fig. 5. Reachability of GSFord and other systems as a function of time

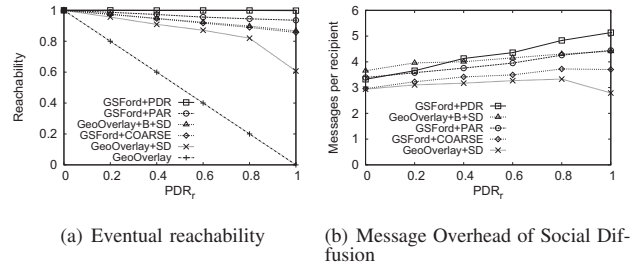


Fig. 6. Performance comparison of GSFord with other systems

from a network emulator, Modelnet[3] with a network topology generated by Inet[4] to accurately mimic Internet scale parameters. On the other hand, since the order of latency of the social diffusion over a phone or an email communication channel is substantially longer than the latency of the dissemination over the delivery overlay of GSFord, we separately modeled the social diffusion process using an Independent Cascade (IC) model[16], [17]. To take into account the disparate levels of interest, social entities may have, to specific messages and the likelihood of communication between neighbors, we enhanced the IC model. Specifically, we exploited multiple probabilities based on the relationship of neighbors and their interest in forwarding the messages. We also enhanced the IC model to consider the latency of delivering messages over multiple communication channels such as an Email channel and a Phone channel.

More detail evaluation settings are found in [20].

B. Reachability of GSFord

We first evaluate the reachability of GSFord to reach the geographically correlated target recipients, TR_g , of a random PAR under a random PDR - the damage is caused by a geographically correlated regional failure. The reachability is defined as the ratio of number of social entities in TR_g receiving the message to the size of TR_g . Note that we do not consider the social entities outside PAR for calculating the reachability, because they most likely receive the message reliably via the geo-aware delivery overlay. We set PAR to be a 16K by 16K meter rectangle centered at a random coordinate inside the global target region, GTG . Experimental results indicate that the reachability of uniform and non-uniform user distributions is very similar; we show the results of the uniform user distribution to conserve space.

We compared GSFord variations under different combinations of geo-aware overlays and social diffusion. *GeoOverlay* models geo-aware overlays[13]; we use the delivery overlay of GSFord as a representative. *GeoOverlay+SD* refers to the case where the delivery overlay is combined with a basic social diffusion process (i.e. without the information overlay). We also compare the value of broadcasting to the entire tree - *GeoOverlay+B+SD* represents the case where

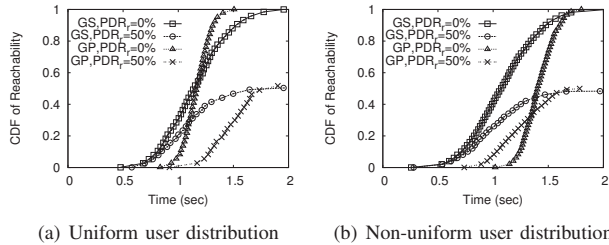


Fig. 7. Comparison of reachability of our RRtree-based delivery overlay of GSFord (GS) with GeoPeer (GP)

a basic social diffusion follows a broadcasting step, where a message is broadcasted over the delivery overlay to the entire *GTG*. In general, such broadcast can be realized using any application layer multicast as a broadcast method[5], [9]. GSFord variations are *GSFord+COARSE*, *GSFord+PAR*, and *GSFord+PDR* which are using coarse, PAR-targeted, and PDR-targeted customization method, respectively.

In Figure 5, we show the reachability and the delay of the event notification process of different systems with different sizes of regional failures. We observe that social diffusion can aid the dissemination of a message significantly when the geo-aware overlay is subject to a regional failure, i.e. where reachability of *GeoOverlay* is bounded by $1 - PDR_r$. While social diffusion can improve message reachability, GSFord can achieve much faster dissemination with even better reachability by leveraging the social information of the target recipients in the *PAR* retrieved from the information overlay. GSFord (with targeted customization with *PAR/PDR*) achieves higher reachability than *GeoOverlay+B+SD*, and reach over 0.8 of target recipients within around 30 minutes after the initial message dissemination, when $PDR_r = 0.8$.

Fig. 6 presents the eventual reachability and the average number of messages that a recipient inside the *PAR* received through the social diffusion process. We observed that leaning to the basic social diffusion is not a good idea to increase the eventual reachability. Employing broadcasting technique (*GeoOverlay+B+SD*) expedites the social diffusion process. However, it achieves only similar reachability of *GSFord+COARSE* despite spending more overhead. On the other hand, GSFord incurs more overhead of social diffusion adaptively to PDR_r and achieves higher reachability. Especially, *GSFord+PDR* achieves almost 1 of reachability under any PDR_r by aggressively encouraging social diffusion process.

C. Performance of RRtree-based Geo-aware Overlay

In this section, we compare the performance of the RRtree-based geo-aware delivery overlay of GSFord with *GeoPeer*, a Delaunay triangulation based geo-aware overlay[13]. Figure 7 shows the reachability of the delivery overlay of GSFord and *GeoPeer* as a function of time under different user distributions. We also consider the impact of the geographical failure ($PDR_r = 0.5$). We observed that the performance of the delivery overlay of GSFord is stable under different user distributions. On the other hand, *GeoPeer* takes higher average delivery times if users are distributed non-uniformly. We also note that the delivery overlay of GSFord can deliver messages faster than *GeoPeer* in the presence of a geographical regional failure. That is, the RRtree based geo-aware overlay can support efficient and reliable geographical regional multicasting with both unexpected user distributions and random geographically correlated regional failures. The main reason is that the RRtree based overlay dynamically adjusts both

RRtree and RHT of a PNode based on the population of the region corresponding to the PNode.

VI. CONCLUDING REMARKS

This paper considers the event notification in extreme situations and presents a geo-social notification system, GSFord, where information needs are strongly correlated to both the geographical location of events as well as the social relationships of people. We presented the design of the GSFord system using a reliable geo-aware overlay structure, RRtree over which we build a reliable multicasting protocol. Furthermore, GSFord exploits a social diffusion process to improve/maximize notification coverage and reliability of notification. Reliable storage of social information at trusted points enables GSFord social entities to forward the alert messages to their social friends inside *PAR/PDR*, even under extreme geographically correlated regional failures. Our results indicate that even under 80% infrastructure damage, 90% recipients are reached within 30 minutes via GSFord.

A more detailed analysis of the privacy/security concerns is a topic of further work. Another natural extension of our work is to adapt GSFord to support rapid, reliable dissemination over wireless mobile devices. Our future plans are to study moving failure scenarios, e.g. tornados, in this environment more deeply and design the system for addressing timeliness/reliability for information dissemination using mixed wired/wireless networks.

VII. ACKNOWLEDGEMENTS

This work was funded in part by the NSF Awards Numbers 1057928, 1059436 and 1143705.

REFERENCES

- [1] <http://www.nbclosangeles.com/news/local/Massive-Power-Outage-Hits-San-Diego-OC-129498468.html>.
- [2] <http://www.city-data.com/city/California.html>.
- [3] Modelnet : <http://issg.cs.duke.edu/modelnet.html>.
- [4] Inet : <http://topology.eecs.umich.edu/inet/>.
- [5] K. Kim, S. Mehrotra, and N. Venkatasubramanian, "FaReCast: Fast, Reliable Application Layer Multicast for Flash Dissemination," in Proc ACM Middleware 2010.
- [6] M. Deshpande, B. Xing, I. Lazardis, B. Hore, N. Venkatasubramanian, and S. Mehrotra, "CREW: A Gossip-based Flash-Dissemination System," in Proc. ICDCS 2006.
- [7] K. Kim, and N. Venkatasubramanian, "Assessing the impact of geographically correlated failures on overlay-based data dissemination," in Proc. IEEE Globecom 2010.
- [8] S. Neumayer and E. Modiano, "Network Reliability with Geographically Correlated Failures," in Proc. Inforcom 2010, 2010.
- [9] M. Ferreira, J. Leitao, and L. Rodrigues, "Thicket: A Protocol for Building and Maintaining Multiple Trees in a P2P Overlay," in SRDS'10.
- [10] A. Rowstron, A. Kermarrec, M. Castro and P. Druschel, "Scribe: The design of a large-scale event notification infrastructure," in Networked Group Communication, 2001.
- [11] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Application-Level Multicast using Content-Addressable Networks," In Proc. NGC, 2001.
- [12] K. Seada, and A. Helmy, "Geographic rendezvous-based architectures for emergency data dissemination," In journal of WIRELESS COMMUNICATIONS AND MOBILE COMPUTING 2008.
- [13] F. Araujo and L. Rodrigues, "GeoPeer: A Location-Aware Peer-to-Peer System," In Proc. NCA 2004.
- [14] J. Zhang, G. Zhang, and L. Liu, "GeoGrid : A Scalable Location Service Network," In Proc. ICDCS 2007.
- [15] J. H. Ahnn, Uichin Lee, and H. J. Moon, "GeoServ: A Distributed Urban Sensing Platform," In Proc. CCGrid 2011.
- [16] D. Kempe, J. M. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in Proc. KDD 2003.
- [17] C. Budak, D. Agrawal, and A. E. Abbadi, "Limiting the Spread of Misinformation in Social Networks," in Proc. WWW 2011.
- [18] N. Kourtellis, J. Finnis, P. Anderson, J. Blackburn, C. Borcea and A. Iamnitchi, "Prometheus: User-Controlled P2P Social Data Management for Socially-aware Applications," in Proc. ACM Middleware 2010.
- [19] M. Sirivianos, K. Kim, and X. Yang, "SocialFilter: Introducing Social Trust to Collaborative Spam Mitigation," Infocom'11.
- [20] K. Kim, Y. Zhao, and N. Venkatasubramanian, "GSFord : A Reliable Geo-Social Notification System," http://altair.chonnam.ac.kr/~kbkim/data/gsford_tech_report.pdf