We designed a smartphone-based system that allows people infected with a contagious disease to alert others to whom they may have spread the virus. We address three challenges: finding devices in close encounters with minimal changes to existing infrastructure, ensuring authenticity of alerts, and protecting privacy of all users. Finally, we also consider the challenges of a real-world deployment.

1. INTRODUCTION

Location-based sharing services allow people to connect with others who are near them, or with whom they shared a past encounter. Suppose it were also possible to connect with people who were at the same location but at a different time—we define this scenario as a close encounter, i.e., an incident of spatial and temporal proximity. By detecting close encounters, a person infected with a contagious disease could alert others to whom they may have spread the virus. We designed a smartphone-based system that allows people infected with a contagious virus to send alerts to other users who may have been exposed to the same virus due to a close encounter. We address three challenges: finding devices in close encounters with minimal changes to existing infrastructure, ensuring authenticity of alerts, and protecting privacy of all users. Finally, we also consider the challenges of a real-world deployment.
be exposed to a contagious virus because of their spatio-temporal proximity to the infected user.

One approach for personalization is for the system to provide exposure information to anyone who may have visited the same location as the diagnosed patient. Location-specific alerts, however, may cause people visiting the location to panic and be concerned about visiting the location; in some cases, the location may be unavoidable (such as a classroom). Spatio-temporal personalization is better, as only those users who may have been exposed to the virus will receive personalized information; every other user using the application will receive a generic alert, with contents similar to the campus-wide email.

Although the broader effects of this idea on campus health will need further study via experimental deployment, this paper focuses on the technological foundations of the detection of close encounters. We make four contributions:

- We introduce a new problem: detecting close encounters. Close-encounter detection allows location-based services to also connect people in temporal proximity, in addition to those who are co-located.
- We present the design of a smartphone-based system called ENACT (Encounter-based Architecture for Contact Tracing) to detect close encounters, while providing strong privacy properties.
- We describe protocols for the ENACT architecture that send alerts to the users who were in spatial and temporal proximity of the infected user. We further extend the protocol to reduce the risk of disclosure of user identity or location history.
- We consider the challenges of a real-world deployment of the ENACT protocol and suggest ways to handle the challenges.

2. THE ENACT SYSTEM

Even though all users receive the alert, the system must, with high accuracy, identify those users (recipients) who were near the infected surface during the time the virus was active, and alert them about their exposure to the virus. We begin with a simple design that addresses personalization accuracy, but later, we extend the design to also provide user privacy and alert authenticity.

2.1 System model

A simple approach to identify the users who were in a close encounter with the patient could be to compare the location traces of the users with those of the patient. Instead of storing users’ actual location traces, ENACT relies on “event tags” – the tag encapsulates both the spatial and temporal location of the user, i.e., where the user was located and when.

The ENACT system comprises five main components: users, smartphones, transmitters such as Wi-Fi access points, health providers, and the ENACT server. We make the following assumptions about each type of component.

Users. Users are people who have installed the ENACT mobile application on their smartphones and carry their smartphones at nearly all times; a study conducted in 2013 with 7,446 participants showed that 79% of smartphone users have their phone on or near them for all but two hours of their waking day and 63% keep it with them for all but one hour [12].

Smartphones. Smartphones have wireless capabilities to receive packets from nearby transmitters and connect to the Internet. On receiving packets, the application stores a tag that contains both the identity of the transmitter and the time when transmitter was observed by the smartphone. The application also stores, in the tag, the RSSI value, which is a number representing the strength of the transmitter signal.

Health Providers. Only an authorized health provider can diagnose a patient to be sick and contagious, collect tags from a patient’s smartphone, and send alerts about the disease; we discuss later how the ENACT can verify authenticity of alerts.

Wi-Fi access points (AP). In the simple protocol design, the tags contain the MAC addresses of APs that the user’s smartphone encounters (instead of the actual locations). If two users are at the same location, their smartphones will observe at least a threshold number of common APs; this fact is the basis of the tag matching for detecting close encounters.

ENACT Server. The server receives alerts and forwards them to all users.

2.2 Event tags

The smartphone application stores event tags, which contain the identity of the Wi-Fi access points (APs) it observes as well as the time duration during which they were observed. How can we use this information to determine if two users shared a close encounter?

Suppose the users shared an encounter. One way of determining if two devices are spatially close is by finding their individual locations and then computing the distance between them [14].

Proximity-based localization techniques rely on the assumption that a smartphone is approximately co-located with the wireless device from which it receives the strongest signal. Prior work has relied on the correlation of the wireless (RF) signals obtained by two devices in close proximity to each other to determine whether they were spatially co-located [14, 15, 20]. ENACT relies on a similar hypothesis that co-located devices observe a threshold number of common APs with similar RSSI values, ignoring the vagaries of wireless networks. People’s durations of stay may vary, but they will have both observed the common access points during the time when they both were at the same location, with similar RSSI values [14, 15]. At least one pair of event tags will match as they observed at least one common AP with similar RSSI values, and within the same minute, since they shared an encounter.

Suppose the users shared a close encounter. Since the patient and recipient were at the same location, they would have observed at least a threshold number of common APs with similar RSSI values. But they observe the APs at different times, and these times are also recorded in the tags. The event tags will match only if the common APs were observed within the temporal boundary of the close encounter. If a close encounter was defined by “all users who visited a location within two hours”, then two users’ tags match only if they observed a threshold number of common APs with similar signal strengths within a span of two hours.

For example, at time $T$, $A$ observes APs $a$, $b$, and $c$, and at time $T + 60$, i.e., an hour later, $B$ observes $a$ and $d$. A stores the tags $(T, a, R_a)$, $(T, b, R_b)$, and $(T, c, R_c)$ and $B$ stores $(T', a, R'_a)$, $(T', b, R'_b)$, and $(T', c, R'_c)$, where $T' = T + 60$ and $R_a$ and $R'_a$ indicate $a$’s signal strength as observed by $A$ and $B$ respectively. If a close encounter is defined by a temporal boundary of two hours, $R_a \sim R'_a$, and the threshold number of common APs observed at that location is at least two, $A$ and $B$ will have one set of matching tags, stored at times $T$ and $T'$. If $B$ continues to observe $a$ and $b$ with similar signal strengths until $T' + n$, then there will be $n$ matching tags, indicating that $B$ was exposed to the virus for $n$ minutes.
The ENACT system relies on the smartphones to store the tags until a disease outbreak occurs. Since the incubation period for most contagious diseases is less than a month, it may be sufficient to store tags for a month. However, a user may choose to delete tags at any time, at the cost of losing the opportunity to obtain personalized alerts about virus exposures.

2.3 Communication steps

Now, we present the protocol and explain how the different components of the ENACT system interact with each other.

1. All smartphones store the MAC addresses of the APs they observe, along with the RSSI and the time when the smartphone observes the APs.

2. When a user is diagnosed by the health provider, the health provider generates an alert using the ENACT application using disease and patient-specific parameters and sends it to the patient, via the ENACT app. The patient-specific parameters include how long the patient has been contagious and the time a virus survives on a surface.

3. On obtaining the patient-specific parameters, the patient’s smartphone retrieves all the tags it stored during the time duration when the patient was contagious, attaches it to the alert and sends it to the server.

4. The server forwards the message to all users. On obtaining the alert, the recipient’s application retrieves the relevant tags stored on the user’s smartphone, during the time when the patient was contagious. The application tries to match the tags it retrieved with the tags included in the alert message. If there are n matches, the application generates and provides the recipient a personalized alert that the user may have been exposed to the virus for n minutes. If there are no matches, the application presents the user a general alert from the health provider.

3. EXTENDING ENACT

Since tags contain the MAC addresses of APs, on obtaining the tags, a recipient application can determine which APs a patient saw and when. If a recipient with malicious intent gains access to this information, they may be able to determine the identity of the patient, or reconstruct the location history of the patient. Other ways malicious users could misuse the system is to create panic on campus by sending fake alerts. We extend our simple design to provide better privacy to users, and ensure authenticity of alerts.

3.1 User Privacy

To prevent tampering and unnecessary exposure of location history, the system needs a different location representation that varies across time and space. This representation should not statically refer to the location, or to a proxy for location like the MAC address of an AP, so any compromise of the location history will not disclose actual locations. However, this representation needs to be generated by the AP itself since both the patient and recipient should be able to obtain it. Furthermore, the representation, which we call an ‘event tag’, should vary over time so no adversary can “map” the entire campus and be able to determine users’ past locations. Instead, these tags should be designed solely for determining close encounters. The patient’s smartphone must be able to generate event tags for the duration when the virus is active, for each location they were at when contagious, because the patient may not stay in or near one location for the entire period when she was contagious and may receive only one or a few tags from the same access point when she was at that location.

In the ENACT system, an access point uses hash chaining to generate and broadcast a series of related numbers; it broadcasts a new number every b minutes. For simplicity, assume for now that the hash chain grows infinitely in that it never gets reset. In later sections, we discuss how the ENACT system handles hash-chain reset.

3.2 Alert authenticity

Next, to prevent fake alerts, the system must allow only the smartphone of a patient who is sick and contagious, as diagnosed by an authorized health provider, to send alerts about the disease. We achieve this goal by adding one more component to the system, a PKI server to issue certificates for health providers. When the health providers install the ENACT application on their smartphones, they generate private-public key pairs and register with the PKI server to obtain a certificate for their public key.

When a user is diagnosed by the health provider, the health provider generates an alert, signs the message and sends it to the patient, via the ENACT app. After attaching the tags, the patient’s smartphone sends the new message back to the health provider. The health provider signs the final message that includes the signed alert and list of tags from the patient, and sends the signed message to the ENACT server. The server forwards the signed alert to all the users. On receiving the alert, the user’s smartphone verifies the signature on the alert. If the signature is not valid, the message is discarded. Otherwise the application attempts to find and match tags in the user’s smartphone with the tags in the alert it received.

The updated protocol with the hash-chaining technique and digital signatures works as follows; the protocol steps are shown in Figure 1, using notation from Table 1.

1. Beacon broadcast: The wireless transmitter, e.g., a Wi-Fi access point, sends a number representing location that we refer to as tag t. The smartphone obtains it with signal strength R at time T, and adds to its database a tuple (t, T, R); for simplicity, we exclude R in the diagram and subsequent descriptions. If it receives multiple tags from different access points during the minute T, it stores them as separate tuples,

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>Patient</td>
</tr>
<tr>
<td>r</td>
<td>Recipient</td>
</tr>
<tr>
<td>u</td>
<td>Other users</td>
</tr>
<tr>
<td>HP</td>
<td>Health provider</td>
</tr>
<tr>
<td>B</td>
<td>Wireless transmitter</td>
</tr>
<tr>
<td>n</td>
<td>Number of wireless transmitters at any location</td>
</tr>
<tr>
<td>b</td>
<td>Broadcasting interval</td>
</tr>
<tr>
<td>t</td>
<td>Hash chain value</td>
</tr>
<tr>
<td>T</td>
<td>Timestamp</td>
</tr>
<tr>
<td>C</td>
<td>Contagious period</td>
</tr>
<tr>
<td>τ</td>
<td>Virus survival period</td>
</tr>
<tr>
<td>R</td>
<td>RSSI value</td>
</tr>
<tr>
<td>g(t, T, R)</td>
<td>Set of generated tags based on location patient was at time T when receiving signal of RSSI R</td>
</tr>
<tr>
<td>h(X)</td>
<td>Message digest (hash) of X</td>
</tr>
<tr>
<td>f</td>
<td>Position in hash-chain</td>
</tr>
<tr>
<td>d(T)</td>
<td>Database record for time T</td>
</tr>
<tr>
<td>sig(D, X)</td>
<td>D’s signature on text X</td>
</tr>
<tr>
<td>E(k, X)</td>
<td>Encrypt X using key k</td>
</tr>
</tbody>
</table>

Table 1: Notations used in the ENACT system model.
The health provider’s message also contains a cryptographic tag, and at the end of the broadcasting interval of $B$, it removes any duplicates.

Consider an example scenario to better understand the tag-generation process: suppose the patient is at a location for an extended period of two hours, the broadcasting interval is one minute, and the contagious period is three hours; it uses the tag it received when it reached the location as the base number to generate numbers for the remaining duration of the hash-chain for up to three hours. So the application on the patient’s phone adds to the alert message the 120 tags it generated during the two hours, and generates the rest of the 60 tags by applying the hash function repeatedly, starting with the last tag it received when at the location. Of course, a patient may be contagious for days, and will be at different locations for varying durations at different times during the contagious period. For each tag obtained during the contagious period, the application creates a hash chain of 180 numbers, and removes any duplicates.

The smartphone collects all the tags in a message and encrypts the set with the key sent by the health provider, and sends the encrypted message back to the health provider. The health provider decrypts the message, generates an alert message $A$ and attaches the tags to it. The health provider then signs the alert message and sends it to the server. The server forwards it to all users.

2. Sending alert: When a patient $p$ observes symptoms, she visits the health provider $HP$ and may be diagnosed with a certain disease. The health provider generates a message that contains information pertaining to the patient and her contagious state, including the time period during which she was contagious, $C$. The health provider’s message also contains a cryptographic key $k$, and her signature to ensure the message’s authenticity.

On receiving the message, the patient’s smartphone retrieves all the tags it obtained during the time period $C$, so for all $T \in C$, it finds all the tuples of the form $(t, T)$ and retrieves the tags $t$. For every tag it obtains, the smartphone computes a set of tags that may have been received by people who were at different locations the patient visited during the virus survival time $\tau$ at those locations. This set $g(t, T)$ depends on the tags it obtained at time $T$ and differs for each tag it obtained from different access points; so if the smartphone obtained $n$ tags of the form $t_1, t_2, \ldots, t_n$, at time $T$, it would generate $n$ sets of $g(t, T)$.

Suppose the patient’s smartphone obtains $t_i = h^i(a_i)$ from $B_i$ at time $T$. The smartphone generates the set of tags $g(t_i, T)$ by generating $\frac{\tau}{b}$ numbers forward in the hash chain, starting from the most recent number $t_i$; $\frac{\tau}{b}$ denotes the number of tags the access point will send during the virus survival period of $\tau$ minutes. Then $g(t_i, T) = \{h^j(t_i) | 0 \leq j < \frac{\tau}{b}\}$. The application eliminates duplicate tags among the generated sets; duplication may happen in tag generation by the application if the user stayed at the location for longer than one broadcasting interval and received multiple tags from the same access point.

Consider an example scenario to better understand the tag-generation process: suppose the patient is at a location for an extended period of two hours, the broadcasting interval is one minute, and the contagious period is three hours; it uses the tag it received when it reached the location as the base number to generate numbers for the remaining duration of the hash-chain for up to three hours. So the application on the patient’s phone adds to the alert message the 120 tags it received during the two hours, and generates the rest of the 60 tags by applying the hash function repeatedly, starting with the last tag it received when at the location. Of course, a patient may be contagious for days, and will be at different locations for varying durations at different times during the contagious period. For each tag obtained during the contagious period, the application creates a hash chain of 180 numbers, and removes any duplicates.

The smartphone collects all the tags in a message and encrypts the set with the key sent by the health provider, and sends the encrypted message back to the health provider. The health provider decrypts the message, generates an alert message $A$ and attaches the tags to it. The health provider then signs the alert message and sends it to the server. The server forwards it to all users.
3. Viewing alert: On receiving the alert message, the smartphone first verifies the signature of the health provider. If the signature is invalid, it discards the message. Otherwise, it retrieves all the tags and attempts to match a threshold number of tags for one pair of time values within the temporal boundary of the close encounter.

Considering only one AP for simplicity, tag matching is based on the following concept: Suppose a user arrived near the access point at a time later than when the patient first observed the same access point. The user receives $h^y(t_i)$ from the same access point. If $y < t_i$, the application will be able to find a pair of values generated from the same hash chain, and hence one matching tag. As shown in Figure 1, if there is at least one such matching tag, the application presents a personalized alert $P$ that informs the user that they were exposed to the virus for at least one minute; i.e., the time of exposure may be computed based on the number of matching tags. If no such match is found, the application shows the user the generic alert generated by the health provider, $A$.

By using the extended protocol, the patient achieves better privacy since their location history is encapsulated in a list of hashes whereas in the basic protocol the MAC addresses of all APs observed by the patient are included in the tags stored on their phone and sent to all users. Our basic protocol can work without modification to any existing network infrastructure since the tags contain MAC addresses; the only caveat to the basic protocol is the exposure of the patient’s location history in the form of list of MAC addresses and time when the MAC addresses were observed. (The non-patient’s location history data is never revealed.)

An adversary could potentially set up fake APs to send hash values, set up listening devices to obtain hash values from nearby APs, or install malware on all phones on campus to obtain hash values, set up listening devices to obtain hash values from nearby APs to generate all the future values to be sent by the AP. However, if the APs occasionally reset their hash chains by picking a new random number $a$ as the seed for a new hash chain, how can we support patients and recipients when the virus survival period near the AP spans a reset moment? A recipient who collected hash values from an AP after the reset moment will not be able to learn about virus exposure from a patient who left the location before the reset occurred, as the patient will only be able to generate hash values in the previous hash chain before the reset and will not have any hash values in common with the recipient. One option is to reset hash chains daily at a time when we expect few close encounters to occur, for example, at 4am.

5. RELATED WORK

Several smartphone applications have been developed to assist contact tracers and responders for monitoring Ebola outbreak through location traces [24], screening for tuberculosis and malaria in Botswana [18], and using text messages [22]. Researchers have developed models to predict the spread of diseases using diagnosis [21] and social network data [26]. ENACT can be used to alert actual contacts, including strangers, based on real encounters and close encounters, without any involvement of contact tracers.

Researchers have used smartphones to help connect strangers who shared an encounter [16, 19], however, ENACT protects the identities of the patient and recipient from each other and the server. Finally, our work is complementary to techniques used for proximity testing [17, 23, 25].

6. SUMMARY

We introduced a new problem: detecting close encounters. We designed the ENACT system that allows people infected with a contagious virus to send alerts to other users who may have been exposed to the same virus due to a close encounter. We presented two designs – a simpler version that focused on personalization accuracy and an extended version for ensuring authenticity of alerts and providing better privacy to the users. Finally, we explored different considerations when implementing the system.

Acknowledgements

This research results from a research program at the Institute for Security, Technology, and Society at Dartmouth College, supported by the National Science Foundation under award number CNS-1329686. We thank Xiaohui Liang, Timothy Lahey, Ann Bracken, Virginia Brack, our colleagues in the THaW project and the anonymous reviewers for their valuable feedback.

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the sponsors.

7. REFERENCES


[12] 79% of people 18-44 have their smartphones with them 22 hours a day [study]. Online at http://www.adweek.com/socialtimes/smartphones/480485 Last accessed April 15, 2017.


