GeoPal: Friend Spam Detection in Social Networks with Private Location Proofs

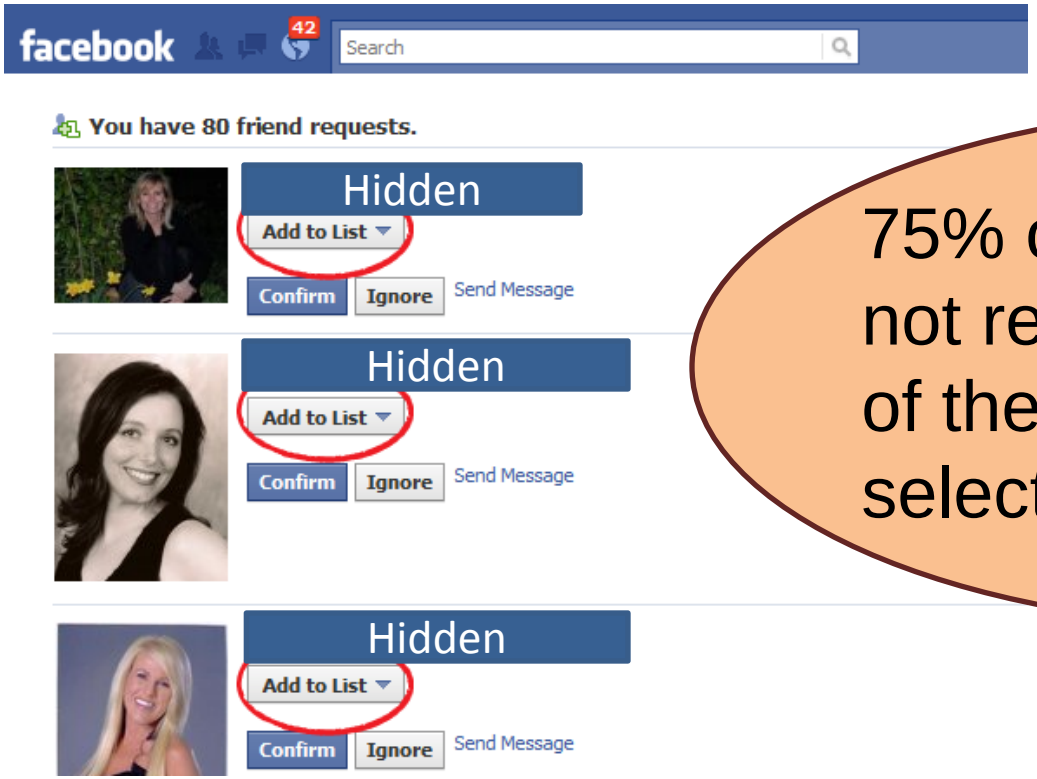
Bogdan Carbunar, Mizanur Rahman, Mozhgan Azimpourkivi, Debra Davis

Florida International University

carbunar@cs.fiu.edu

Social Network Friend Spam

Friend invitations from people you don't know



75% of 68 participants did not remember at least one of their 20 randomly selected friends

Friend Spam Consequences

Attackers can

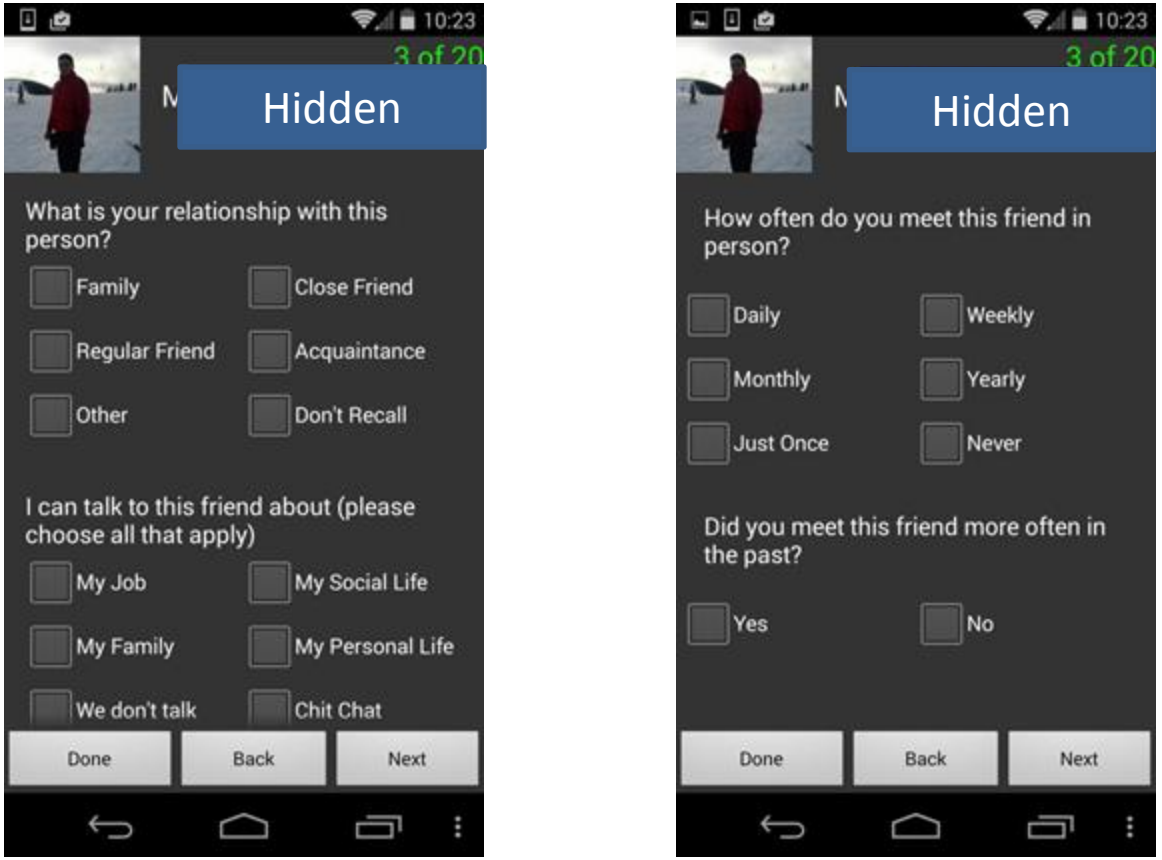
- collect private information from victims
  - profiles, locations visited, friend lists
- spear phishing attacks
- malware dissemination

Assumptions

1. People trust more the friends whom they have met or are meeting more frequently in person
2. Hard to guess locations frequented by the victim
3. Hard to create during the attack a history of co-locations with the victim

Trust vs. Co-Location

People trust more the friends whom they have met or are meeting more frequently in person

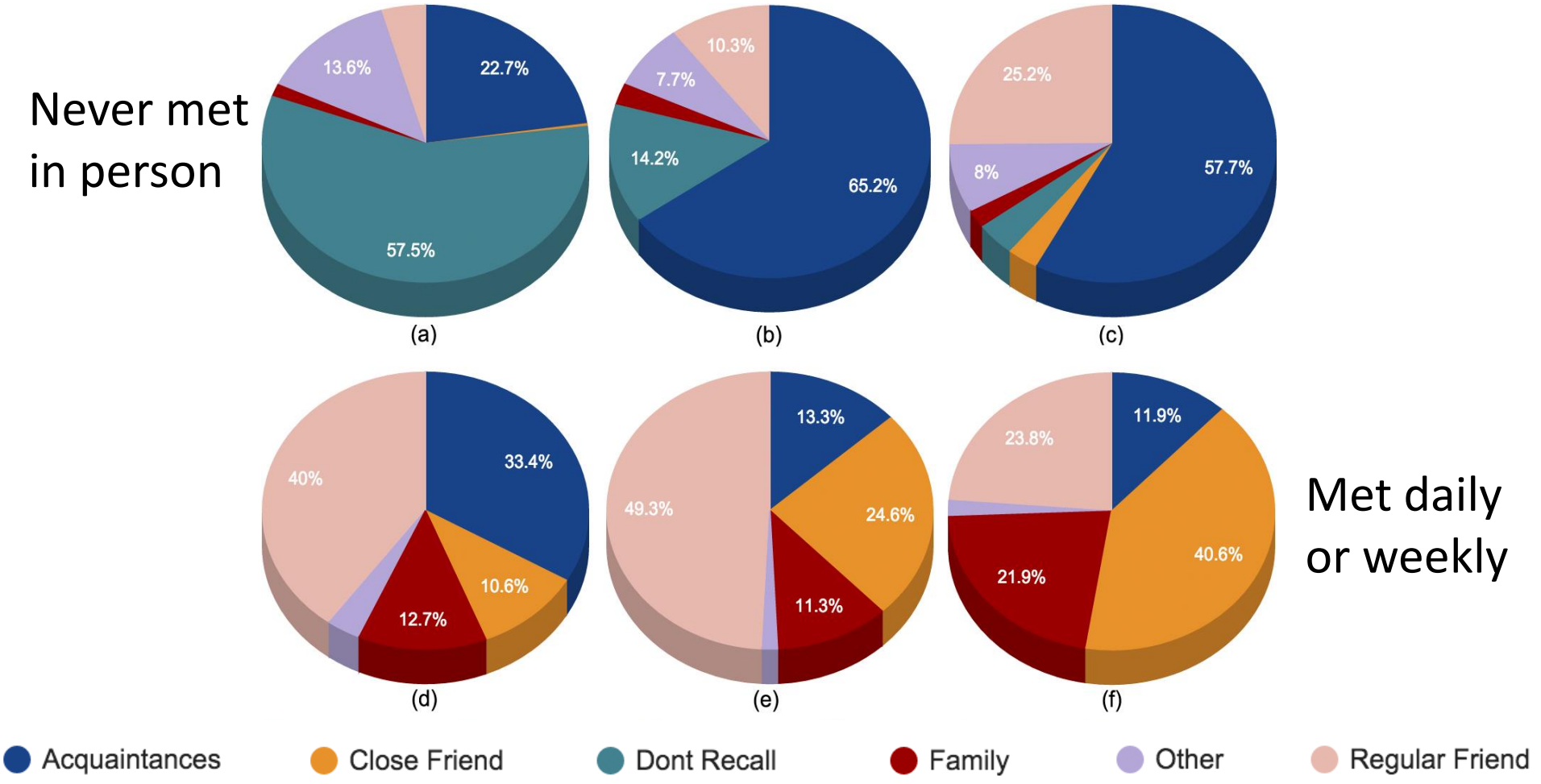


GP.Quest: Mobile App Questionnaire

Location vs. Friend Relationship Quality (Facebook)

- 68 participants (18-50 years old, 57 male/11 female)

Never met in person

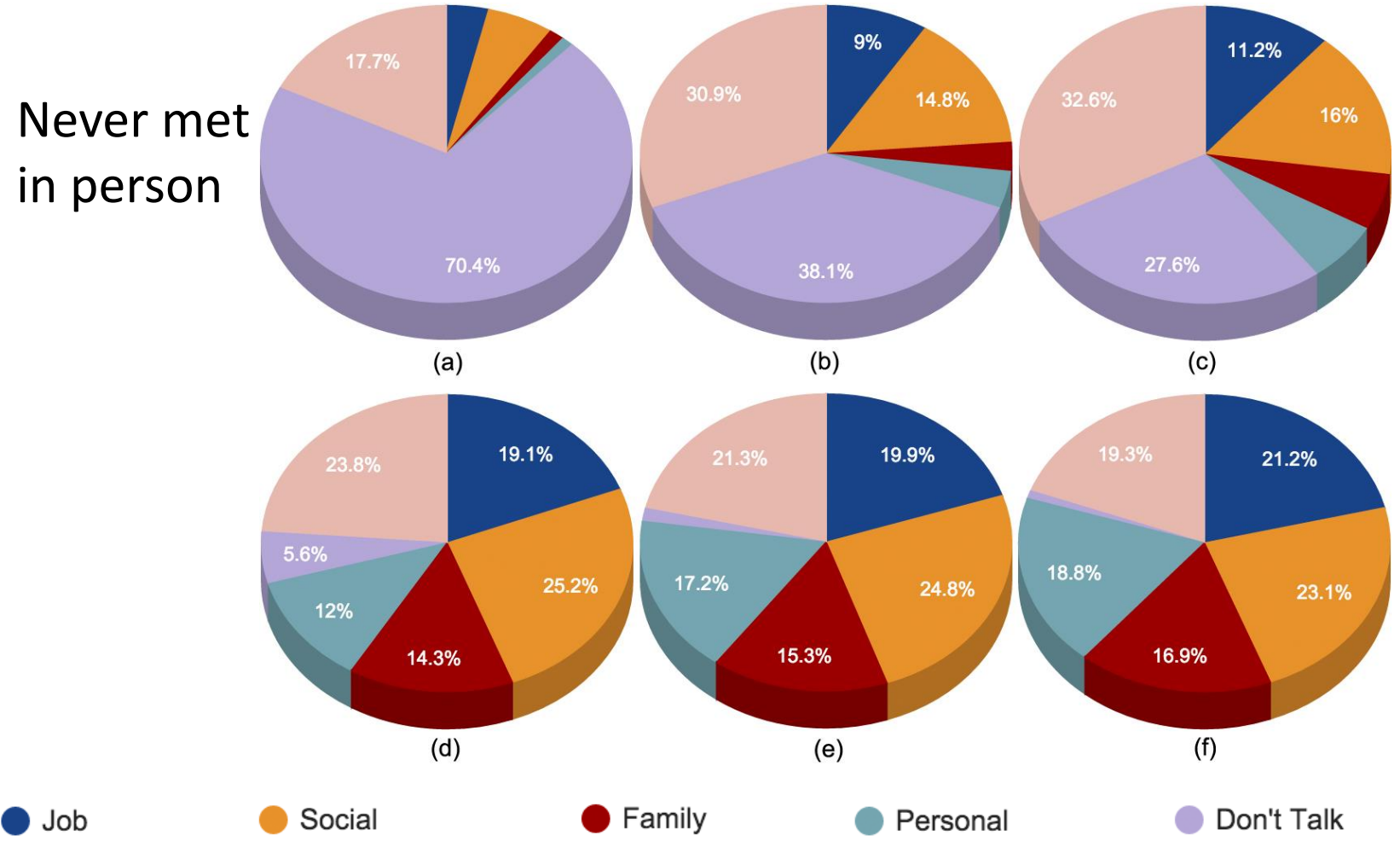


Met daily or weekly

Location vs. Discussion Topics (Facebook)

- 68 participants (18-50 years old, 57 male/11 female)

Never met in person



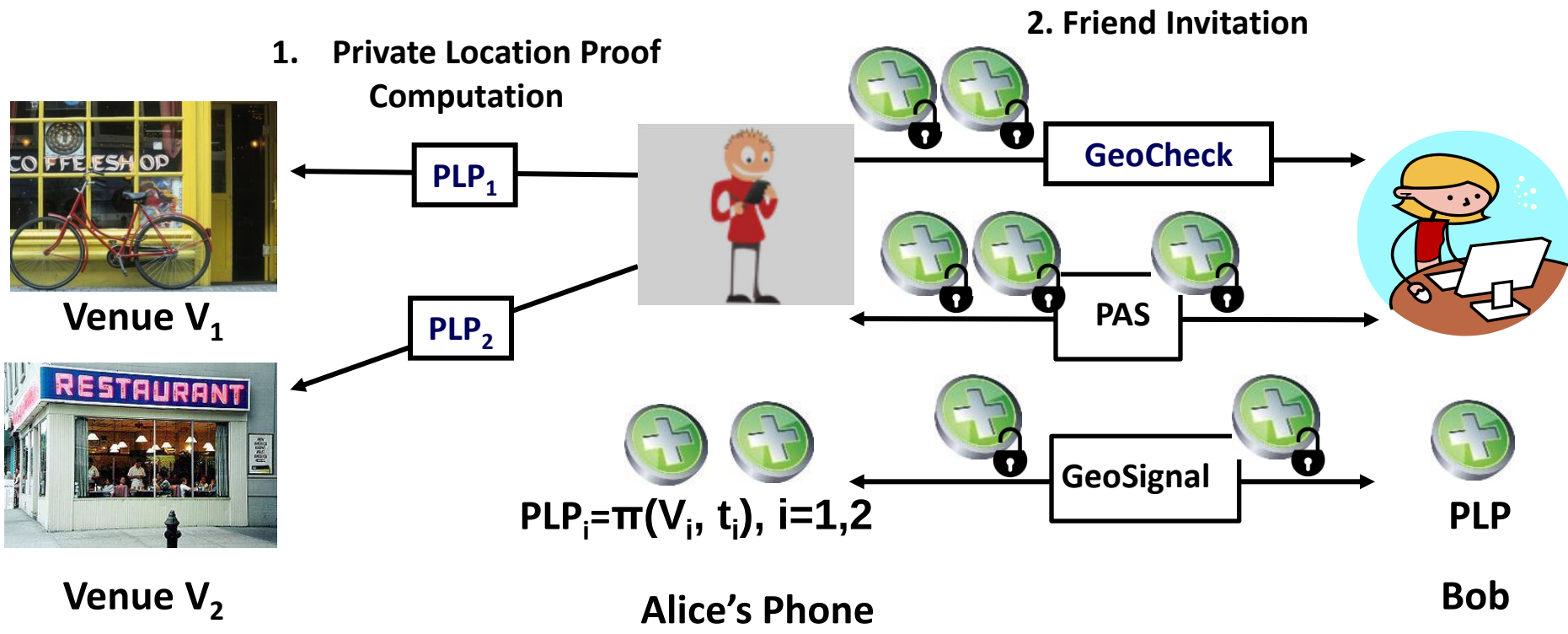
Met daily or weekly

GeoPal

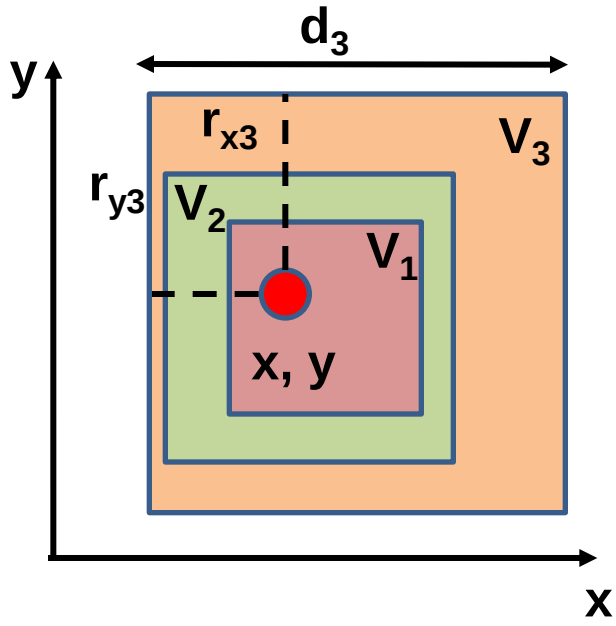
1. People trust more the friends whom they have met or are meeting more frequently in person
2. It is hard to guess locations frequented by the victim
3. It is hard to create during the attack a history of co-locations with the victim

- ❑ Mobile app that records locations visited by user
- ❑ Use location history to establish trust with friends
  - with privacy

GeoPal: Friend Spam Detection Framework

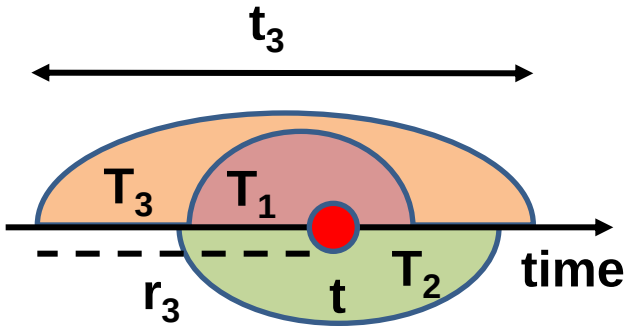


Confusion Zones



$V_3 = [(x-r_{x3}, y+r_{y3}), (x+d-r_{x3}, y-d+r_{y3})]$

Spatial



$T_3 = [t-r_3, t+t_3-r_3]$

Temporal

Presence Tokens

Location V

$Tk_{V,e}$

facebook

Social Network

Social network divides

- ❑ Space at granularity of venues
- ❑ Time at granularity of “epochs” (e.g., 10 min long)

Private Location Proofs

Two users are *fuzzy co-located* when present in the same confusion zone (spatial & temporal)

Private Location Proof

venue & time

Presence token

$$\pi(V,t) = (E_k(Id), V, t, e, Tk_{V,e}, K_{Vi}, K_{Ti}, \overline{V}, \overline{T}, E_V, E_T, \sigma)$$

client pseudonym

key material

obfuscated confusion zones

signature

PLP Construction

"Alice" preserves anonymity!

Alice

1 *E(Id(A)), Time t, Location V*

$(E_k(Id), V, t, e, Tk_{V,e}, K_{Vi}, K_{Ti}, \overline{V}, \overline{T}, E_V, E_T, \sigma)$

Social Network

Location V

2 $\overline{V} = \{V_1, .., V_g\}, \overline{T} = \{T_1, .., T_g\}$ Generate confusion zones

3 $K_{Vi}, K_{Ti}\ i=1..g$ Generate confusion keys

4 $E_V = \{E(K_{Vi}, V_i) \mid i = 1..g\}, E_T = \{E(K_{Ti}, T_i) \mid i = 1..g\}$ Encrypt confusion zones

5 $\sigma = S_{GSN}(E(Id(A)), E_V, E_T)$ Sign location proof

PLP Based User Trust Establishment

GeoPal uses the PLP history to establish trust

- GeoCheck: prove past presence at profile locations
- PFAS: Privately infer co-location affinity with other users
  - How many times the two users *have been* co-located
- GeoSignal: Privately infer present co-location events

GeoCheck: Profile Location Verifications

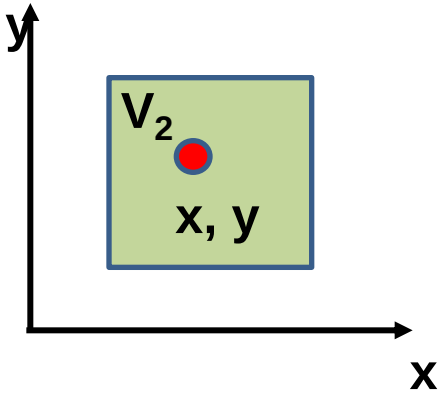
Prove presence around location V around time t

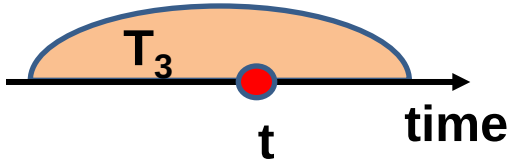
with privacy

$E_V = \{E(K_{Vi}, V_i) \mid i = 1..g\}$, $E_T = \{E(K_{Ti}, T_i) \mid i = 1..g\}$,
$\sigma = S_{GSN}(E(Id(A)), E_V, E_T)$

K_{V2}, K_{T3}

Alice





Bob

Verify **σ**

Decrypt & verify confusion zones

$\pi(V,t) = (E_k(Id), V, t, e, Tk_{V,e}, K_{Vi}, K_{Ti}, \bar{V}, \bar{T}, E_V, E_T, \sigma)$

PFAS: Private Fuzzy Affinity Score

Privately determine co-location frequency of A and B

Compute intersection of sets of tokens (secure multiparty computation)

Alice **Bob**

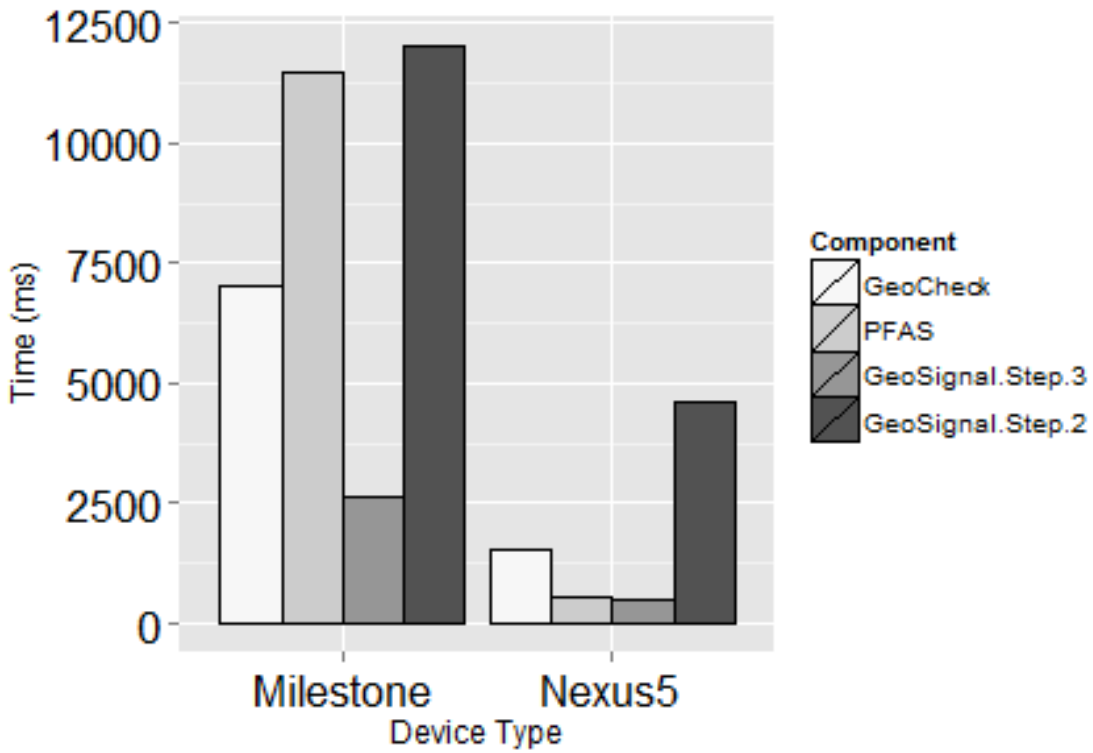
$\pi(V,t) = (E_k(Id), V, t, e, Tk_{V,e}, K_{Vi}, K_{Ti}, \bar{V}, \bar{T}, E_V, E_T, \sigma)$

GeoPal Evaluation

- Motorola Milestone (CPU @ 600 MHz and 256MB RAM)
- Nexus 5 with a Quad-core 2.3 GHz CPU and 2GB RAM

- Industrial grade crypto
  - Signatures: RSA with 2048 bit keys
  - Symmetric encryption: AES
  - Hashes: SHA-512

GeoPal is Practical



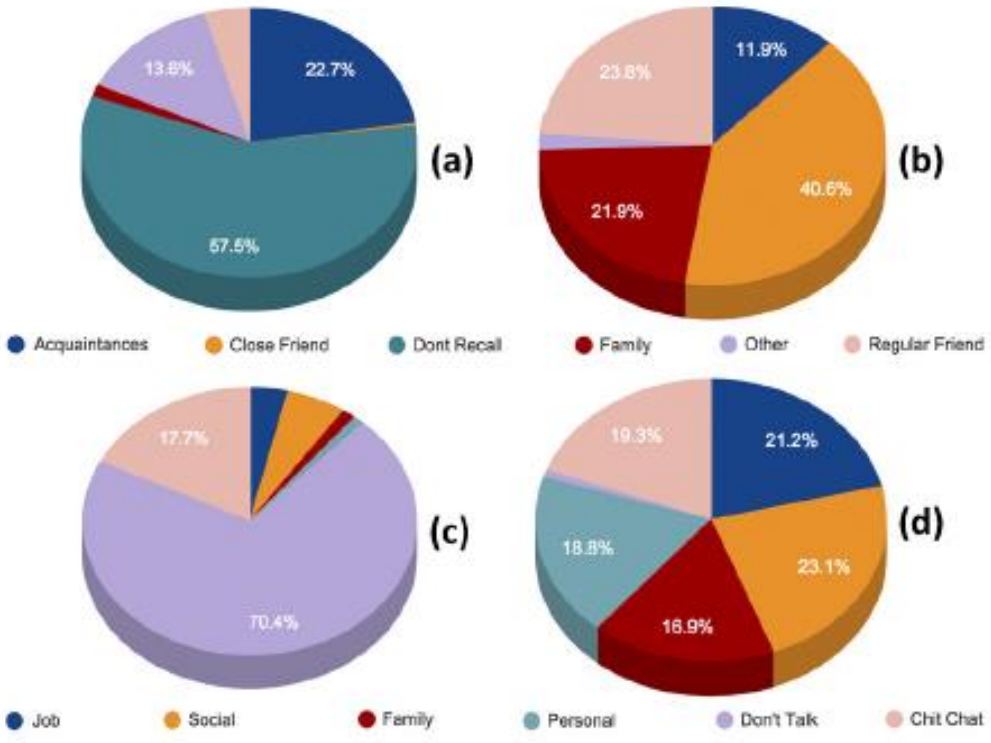
Nexus 5:

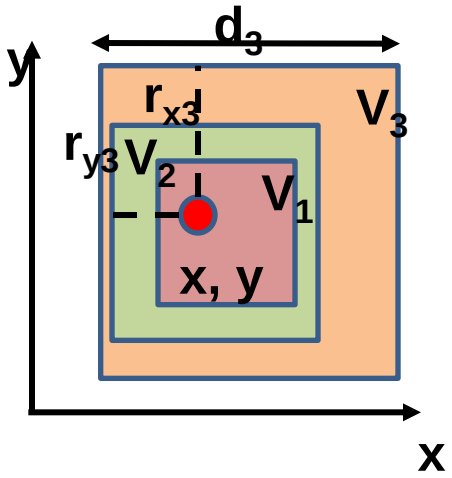
- ❑ 1.5ms to verify a location claim
- ❑ 1s to verify co-location over 20K+ location proofs

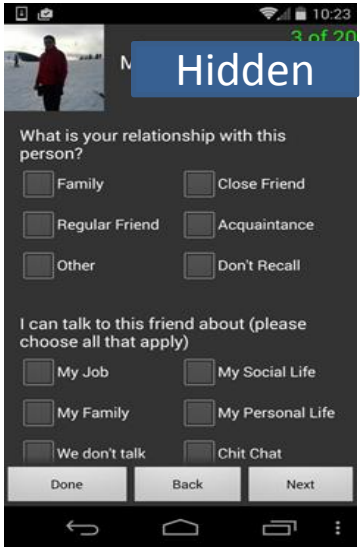
Conclusions

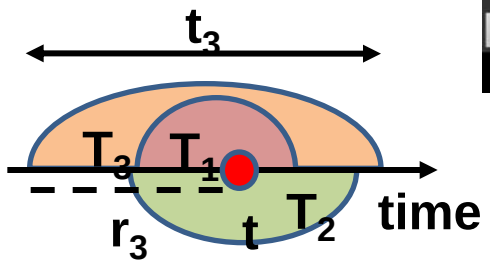
- User study: trust vs. co-location frequency relationship
  - Friend relations stronger with increased co-location
  - More discussion topics with frequently met friends
- GeoPal: seamless, location based friends spam detection
  - Exploit location history to establish trust with friends
- With privacy:
  - Alice learns nothing from Bob
  - Alice controls what she reveals to Bob
  - The social network does not learn Alice's locations

Questions


13.6%
22.7%
57.5%
(a)
23.8%
11.9%
40.6%
21.9%
(b)
Acquaintances
Close Friend
Dont Recall
Family
Other
Regular Friend
17.7%
70.4%
(c)
19.3%
21.2%
18.8%
23.1%
16.9%
(d)
Job
Social
Family
Personal
Don't Talk
Chit Chat


d_3
y
r_{x3}
V_3
r_{y3}
V_2
V_1
x, y
x


10:23
Hidden
What is your relationship with this person?
Family
Close Friend
Regular Friend
Acquaintance
Other
Don't Recall
I can talk to this friend about (please choose all that apply)
My Job
My Social Life
My Family
My Personal Life
We don't talk
Chit Chat
Done
Back
Next


t_3
T_3
T_1
r_3
t
T_2
time