Private GeoSocial Networks

Bogdan Carbunar
Florida International University
carbunar@cs.fiu.edu

Radu Sion
Stony Brook University
sion@cs.stonybrook.edu

Rahul Potharaju
Purdue University
rpothara@purdue.edu

Moussa Ehsan
Stony Brook University
mehsan@cs.stonybrook.edu
Geosocial Networks (GSN)

- Foursquare, Gowalla, Facebook Places, SCVNR, ...

Check-In

Check-In

Check-In
Check-Ins

- Register presence at a venue *now*
- User needs to have GPS equipped phone
  - GPS location is transmitted to GSN
  - Presented with places in vicinity
  - Need to choose one
Badges [Foursquare]

- **Location Badge**
  - \( k \) check-ins
  - At the same venue
  - At different venues
  - Adventurer, Explorer, Superstar, ...

- **Mayor**
  - The most check-ins at a venue in the past two months

- **Multi-Player Badge**
  - Simultaneous check-ins with \( k \) other users
  - Swarm, Super-Swarm, Player Please, ...
User Participation. Why?

- Users receive incentives to participate
  - Lufthansa: ticket discounts for Oktoberfest participants
  - SixFlags: mayor receives season pass
  - Ann Taylor: 25% discount to mayors
  - GAP: 25% off for 2\textsuperscript{nd} check-in
  - Starbucks, Pizza Hut, Burger King
  - ...

Problems

1. Privacy: users provide personal information
   - Name and location traces

2. Cheating
   - Location fraud
   - Badge conditions

- Tension between privacy and cheating
In this talk …

- Data Collection and Statistics
- Secure Location Verification
- Private Badging Protocol
- Evaluation
Foursquare vs. Gowalla

- 780,000 Foursquare users
- 143,000 Gowalla users
  - username, location (home city), friends
  - badges
  - # check-ins, # days out, # things done

Foursquare: Distribution of user home cities in the US

Gowalla: Distribution of user home cities in the US
Foursquare Check-Ins

CDF of days-out, check-ins, and things done per user

- 45% of users have 80-950 check-ins
- Several check-ins per day

Scatterplot of check-ins vs days out
Badges vs. Pins

Foursquare: CDF of friends and badges per user

Gowalla: CDF of pins and check-in per user
In this talk ...

- Data Collection and Statistics
- Secure Location Verification
- Private Badging Protocol
- Evaluation
Why Location Verification?

- Cheating is easy
  - Especially when private
  - GPS Cheat (Android)
  - Location Spoofer (iPhone)

- Users have incentives to cheat
  - Claim fake locations (alibi?)
  - Acquire badges without merit
  - Receive rewards for nothing
Secure Location Verification

- Venue-oriented approach
- Venue owners have the most to lose
  - Reward the wrong customers
- Install small equipment inside venue
- Equipment verifies presence of claimed client
- *One-time investment*
Solution: XACT

- **XACT\(V\):**
  - Generate \(T, \Delta T, S_V(T, \Delta T)\)
  - Encode into Quick Response Code (QRC)
- **User reads and decodes QR code**
- **XACT\(V\) generates new QR code**

**Diagram:***
- **1.** Generate key pair \(pub_V, priv_V\)
- **2.** \(pub_V\)
- **3.** \(T, \Delta T, S_V(T, \Delta T)\)
- **4.** GSN
  - **Alice**
  - Verify signature, expiration
XACT Implementation

- Client: Google Nexus One @ 1GHz
- $\text{XACT}_V$: BeagleBoard Revision C4 @ 720MHz

- Time to generate QR code (BeagleBoard): 50ms
- Time to read and decode QR code (Nexus One): 190ms @ 20cm
In this talk ...

- Data Collection and Statistics
- Secure Location Verification
- Private and Secure Badging Protocol
- Evaluation
Why Privacy?

- GSN provider learns user (location, time) traces
  - User lifestyle
  - Leak, sell to third parties

When in Amsterdam, Alice likes coffee?

- Pseudonymity insufficient [Golle & Partridge; 2009]
  - Anonymous profiles + additional info \Rightarrow Identity
Private Badge Requirements

1. Private check-ins
   - No anonymous profiles – just as damaging

2. Private badge construction

3. Cheating prevention
   - Location
   - Badge conditions

- Need anonymous channels - anonymizer
Approach

1. Check-In
2. Location Proof
3. Token
4. Check-In/Token
5. Check-In/Token
6. ZK Proof of Token Ownership

GSN
Private Mayors

- Mayorship: highest number of check-ins in
  - Past 2 months
  - At most 1 check-in per day

- GSN generates primes $p$ and $q$, modulus $n=pq$
  - $n$ is public
  - $p$ and $q$ are secret

- Once per day, GSN generates random $t$
  - Publish $t^2 \mod n$
  - $t$ is secret
Quadratic Residuosity Assumption

- Given $v$ and $n=pq$
- But not $p$ and $q$
- Hard to say if there exists $x$ such that
  - $x^2 = v \mod n$
Private Mayors

- During check-in, the user:
  - Proves location – over anonymizer
  - Receives square root $t$ of the day’s $t^2 \mod n$
  - Receives blindly signed nonce

- When user has sufficient check-ins ($k$) for mayor
  - Present ZK proof of knowledge of $k$ QR roots +
  - $k$ signed nonces – *do not reuse!*
Zero Knowledge Proof (Part I/III)

- The user has a list
  - L = \{t_1, \ldots, t_k\} of square roots
  - Total is m = 60

- GSN provider has
  - T = \{t_1^2 \mod n, \ldots, t_m^2 \mod n\}
Zero Knowledge Proof (Part II/III)

- User runs the following step s times
  - Generate $Y = \{ y_1, \ldots, y_m \}$ and $z_1, \ldots, z_k$ randomly
  - Set $M = \text{permutation } \pi_1 \{ t_1^2 y_1^2, \ldots, t_m^2 y_m^2 \mod n \}$
  - No need to know square roots of $T$
  - Set $P = \text{permutation } \pi_2 \{ t_1 z_1, \ldots, t_k z_k \mod n \}$
  - Send $M$ and $P$ to GSN

- GSN provider flips coin $b$ and sends to user
  - If $b=0$
    - User sends $Y = \{ y_1, \ldots, y_m \}$ to provider
    - GSN verifies $M$ generated from $T = \{ t_1^2, \ldots, t_m^2 \mod n \}$ and $Y$
Zero Knowledge Proof (Part III/III)

- User runs the following step s times
  - Generate $Y = \{y_1, \ldots, y_m\}$ and $z_1, \ldots, z_k$ randomly
  - Set $M = \text{permutation } \pi_1 \{t_1^2 y_1^2, \ldots, t_m^2 y_m^2 \mod n\}$
  - Set $P = \text{permutation } \pi_2 \{t_1 z_1, \ldots, t_k z_k \mod n\}$
  - Send $M$ and $P$ to GSN

- GSN provider flips coin $b$ and sends to user

- If $b=1$
  - User sends $A = \pi_2 \{a_1 = z_1^{-1} y_1, \ldots, a_k = z_k^{-1} y_k\}$ to GSN
  - GSN verifies that $(t_i z_i a_i)^2 = (t_i z_i z_i^{-1} y_i)^2$ in $M$
ZK Proof Conclusions

- If $b=0$, GSN is convinced that $M$ is built from $T$
- If $b=1$, GSN is convinced that user knows $k$ square roots from $M$
- Thus, with 50% chance, GSN is convinced that user knows $k$ square roots from $T$
- The step is repeated $s$ times
  - Chance of user cheating is $1/2^s$
- GSN only sees one side of the proof
  - Both sides will leak the days when the user checked-in
In this talk ...

- Data Collection and Statistics
- Secure Location Verification
- Private and Secure Badging Protocol
- Evaluation
Experiments

1. Is the server the bottleneck?
   - Check-in
   - Mayor verification
   - 16 quadcore Intel Xeon @2.93GHz and 128GB RAM

2. Can a smartphone support ZK proofs?
   - Google Nexus One @ 1GHz
Mayors

- Dependence on modulus size
  - $k = 30$

- Server
  - Check-in cost independent on bit size – 13,000 per sec
  - Mayor verification - 309 ms for 2048 bit keys

- Client: 7s for mayor generation
Mayors (cont’d)

- Proof generation and verification
- Dependence on number of proof sets
- Dependence on number of check-ins per mayor
  - Linear
  - Server verifies few mayors per second
  - Client takes up to 7s to generate proofs

![Graphs showing time vs. number of proof sets and check-ins](image-url)
Conclusions

- GSN tradeoff between privacy and cheating
- Location verification achieved through venue participation
- Private and secure badge/mayor
- GSN is NOT bottleneck
  - Thousands of check-ins per sec
  - Few mayors verifications/s
- Smartphone supports almost 10 mayor verifications/minute
Questions ?