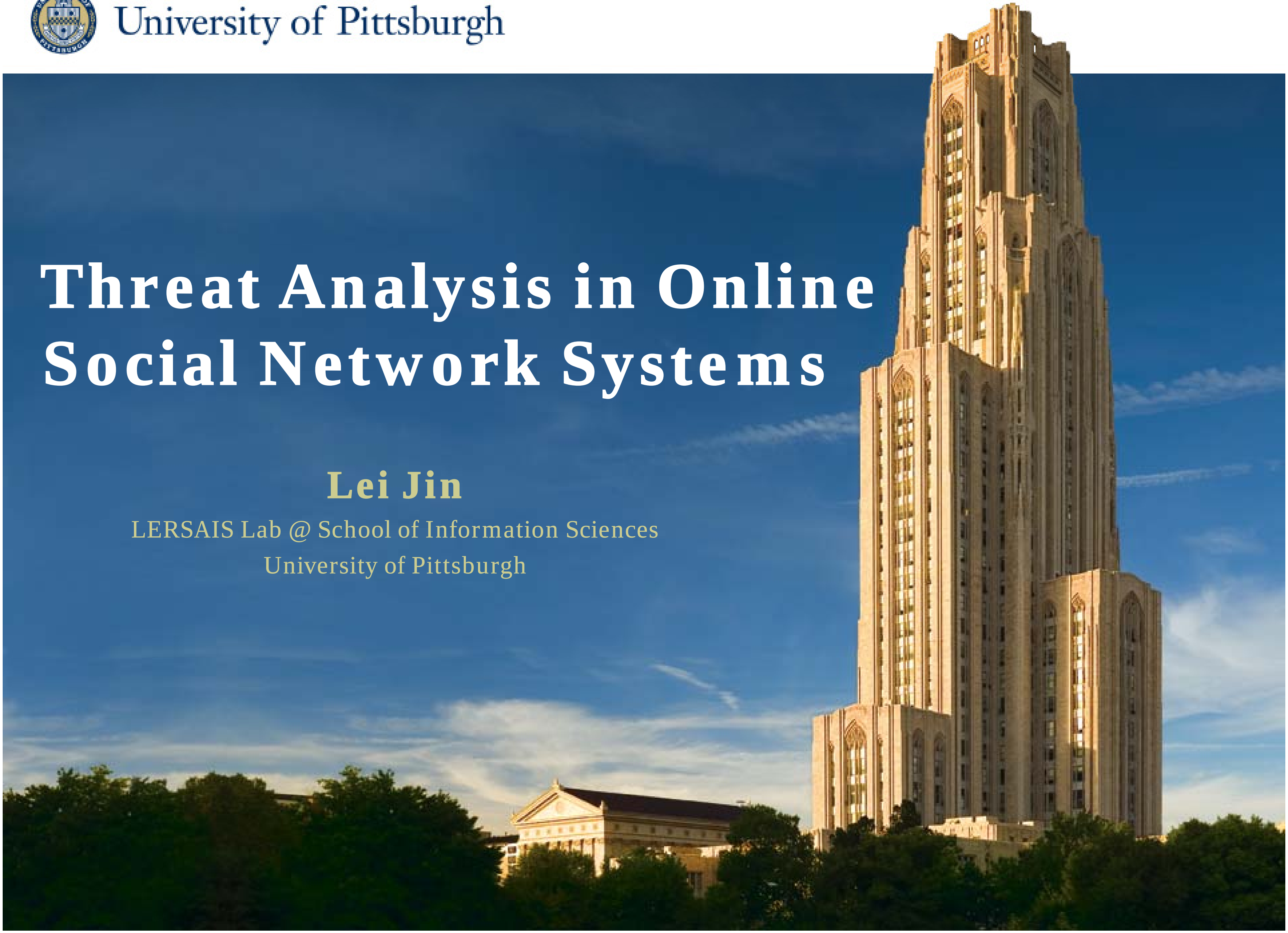
University of Pittsburgh



Threat Analysis in Online Social Network Systems

Lei Jin

LERSAIS Lab @ School of Information Sciences

University of Pittsburgh

Purpose

- Take online social network systems as examples to demonstrate how to conduct a threat analysis in a complex system
- Investigate & analyze various security & privacy issues in the most popular online social network systems, such as Facebook, LinkedIn, Foursquare and Yelp
- Be aware of these problems & know how to mitigate or avoid the potential attacks

Steps of Threat Analysis

- Step1 — Identify Entities, Elements & Mechanisms
- Step2 — Investigate & Test the Mechanisms
- Step3 — Design & Conduct Attacks
- Step 4 — Design & Develop Defense Approaches

Coverage

- Not focus on the traditional problems
  - Authentication
  - Secure Communication
  - Web-based Attacks; E.g., SQL Injection, Cross Site Scripting
- Focus is on the new vulnerabilities that exist in online social networks
  - Traditional online social networks (OSN); E.g., Facebook & LinkedIn
  - Location-based social networks (LBSN); E.g., Foursquare & Yelp

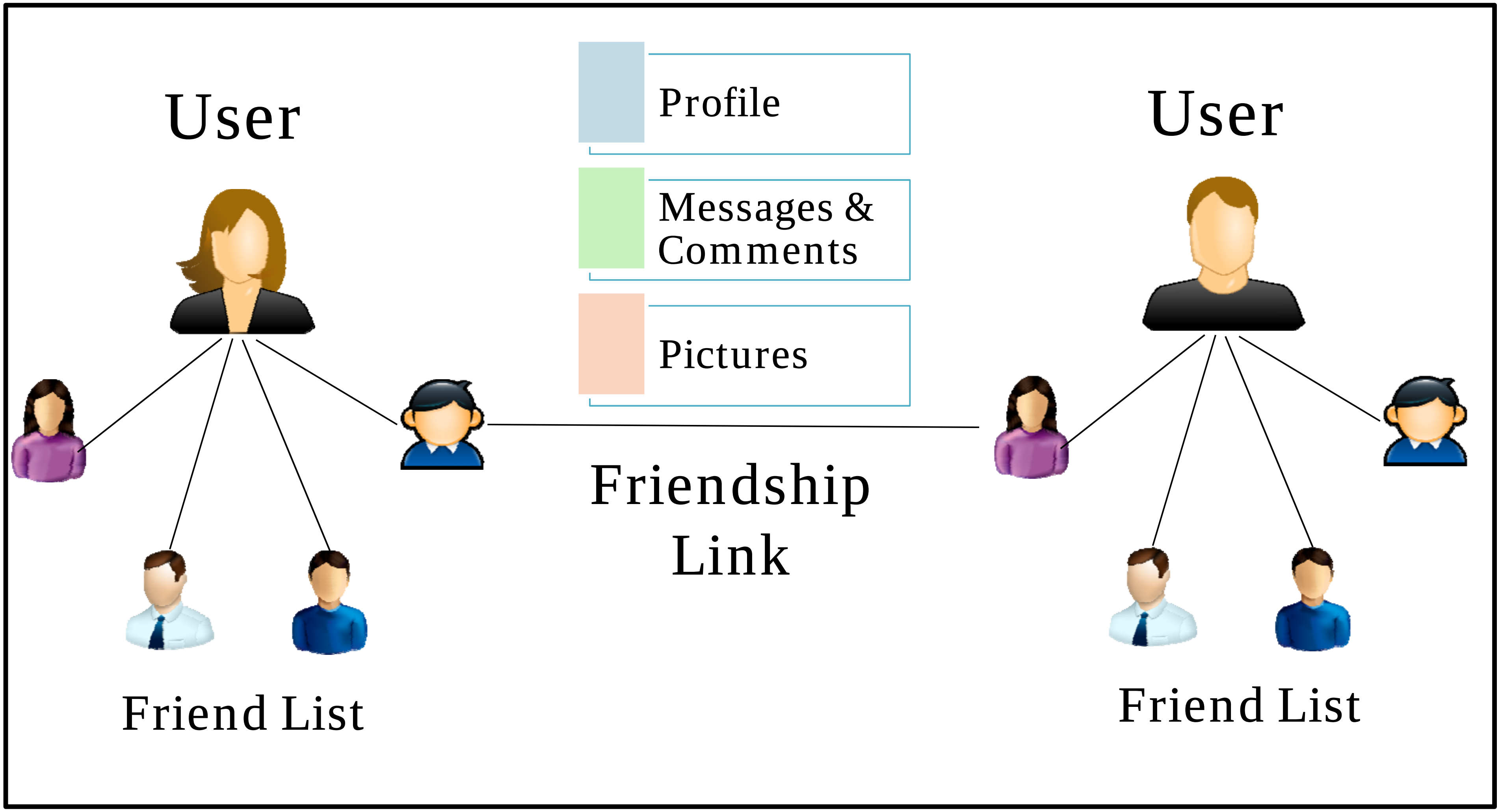
Steps of Threat Analysis

- Step1: **Identify Entities, Elements & Mechanisms**
- Step2: Investigate & Test the Mechanisms
- Step3: Design & Conduct Attacks
- Step 4: Design & Develop Defense Approaches


LinkedIn
skype
LIVEJOURNAL
foursquare
g+
vimeo
Insta
YouTube
facebook
g+
Blogger
my
stagram
twitter
flickr
Bēhanc
tumblr.
Like
yelp

OSN

Friendship/Social Network



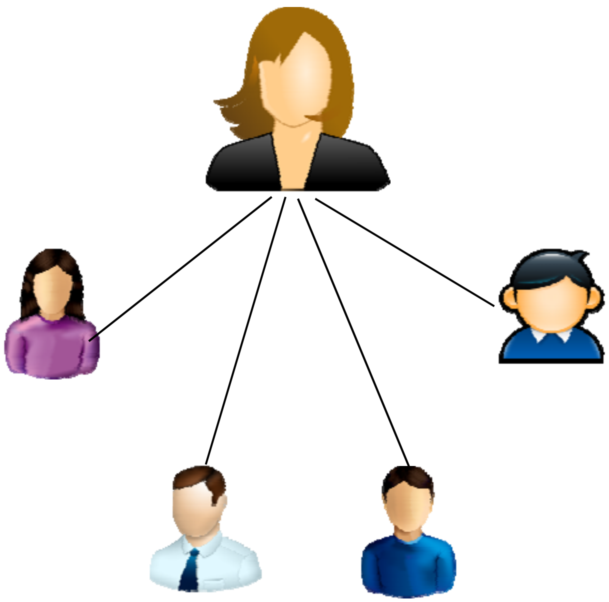
LBSN

Create venues

Explore various places

Check in at venues

User



Social Network

CHECK-IN

(*user*, *venue*, *time*,…)

Venue

VENUE

(*name*, *location*, *category*,…)

Entities, Elements & Mechanisms

- User's Social Network
  - Friends
  - Mutual Friends
  - Recommended Friends
- User's Posts
  - Messages
  - Photos
  - Check-ins (LBSN)

- User Identity / User Profile
  - Attributes
- Venue (LBSN)
  - Attributes

- Mechanisms
  - User Authentication
  - Access Control Mechanisms

Steps of Threat Analysis

- **Step1** — Identify Entities, Elements & Mechanisms
- **Step2** — **Investigate & Test the Mechanisms**
- **Step3** — **Design & Conduct Attacks**

  Attacker!
- **Step 4** — **Design & Develop Defense Approaches**

  Defender!

Investigation of User Identity / User Profiles & Venues

User Identity
- Fake Identity
- Cloned Identity

Identity Clone Attack

User Profile
- Private / Sensitive Attributes
- Exposure of Sensitive Attributes

Attribute Inference Attack

Venue
- Fake Venue
- Fake Promotions & Coupons
- Exposure of Visit Statistics (check-ins)

Venue Attack

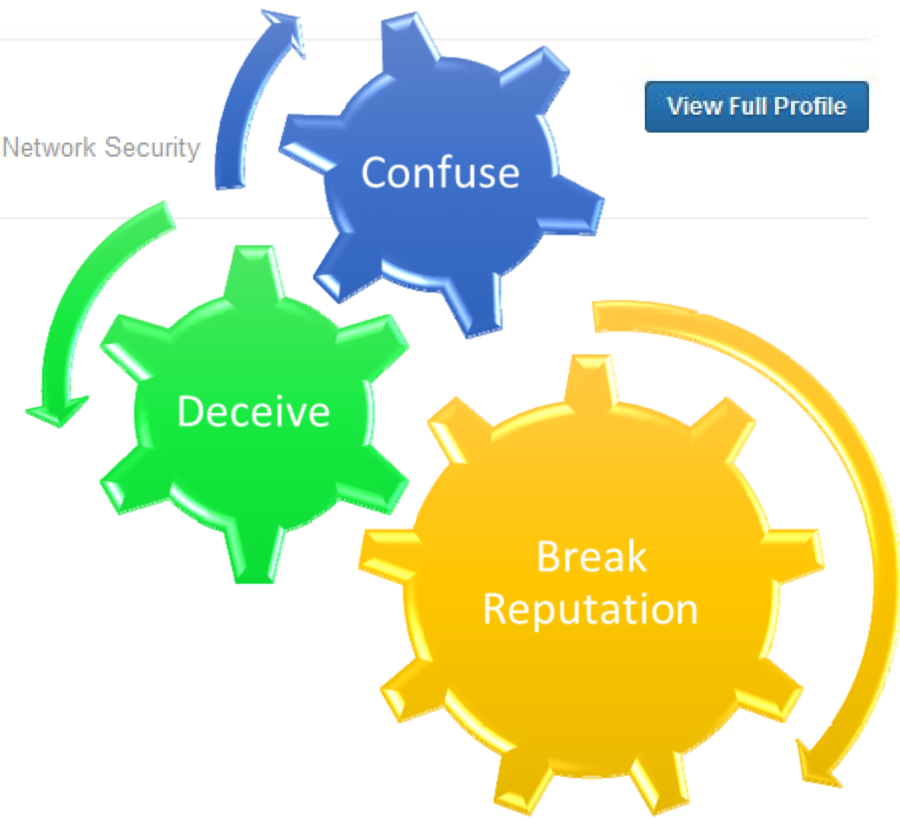
Cloned Identity

Lei Jin
PHD at University of Pittsburgh
Greater Pittsburgh Area | Computer & Network Security
Education University of Pittsburgh, Tsinghua University, Tsinghua University

View Full Profile

Lei Jin
PHD at University of Pittsburgh
Greater Pittsburgh Area | Computer & Network Security

View Full Profile



Identity Clone Attack [6] - Design

- Attributes: name, education, birthday…
- Friend network
  - Friend List (FL): Connected friends of an ID
  - Recommended Friend List (RFL):
    - ✓ Generated by OSN systems (function of ***"People You May Know"*** on Facebook)
    - ✓ Share same RFs
  - Excluded Friend List (EFL):
    - ✓ Social embarrassments
    - ✓ Attackers - try to connect these individuals

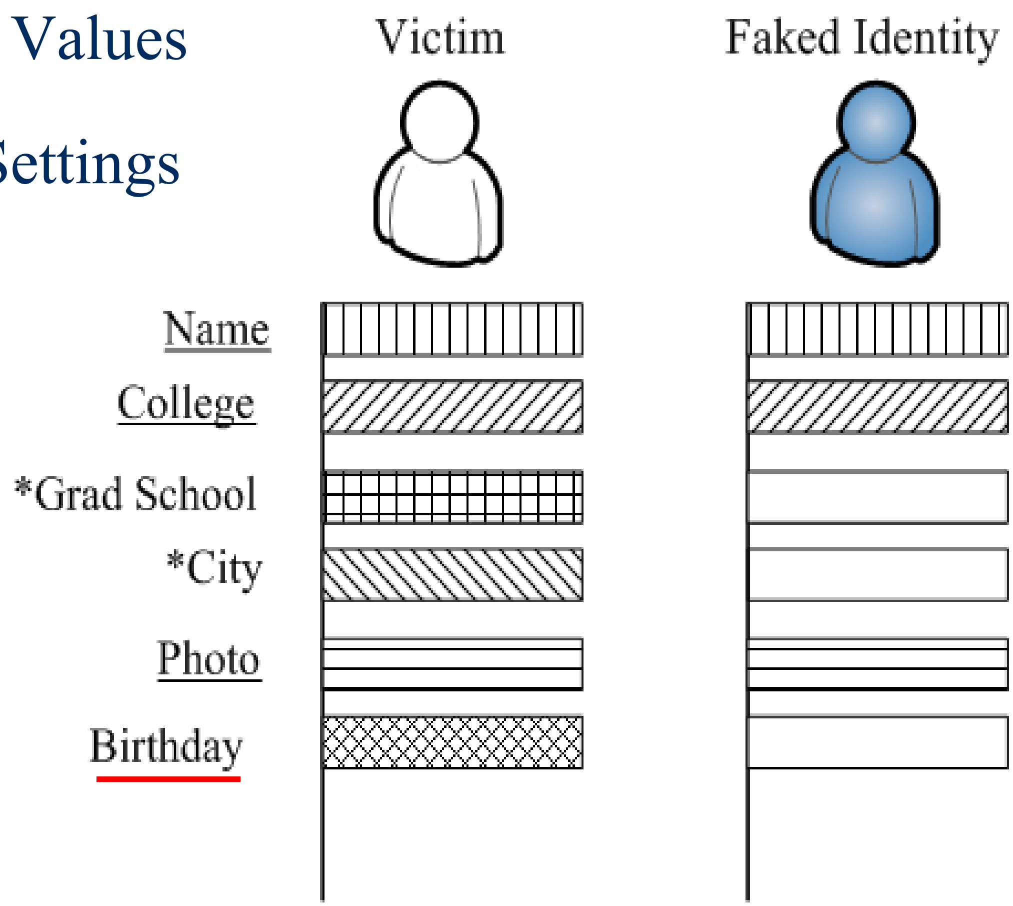
What are the best targets

University of Pittsburgh

Attribute As Target

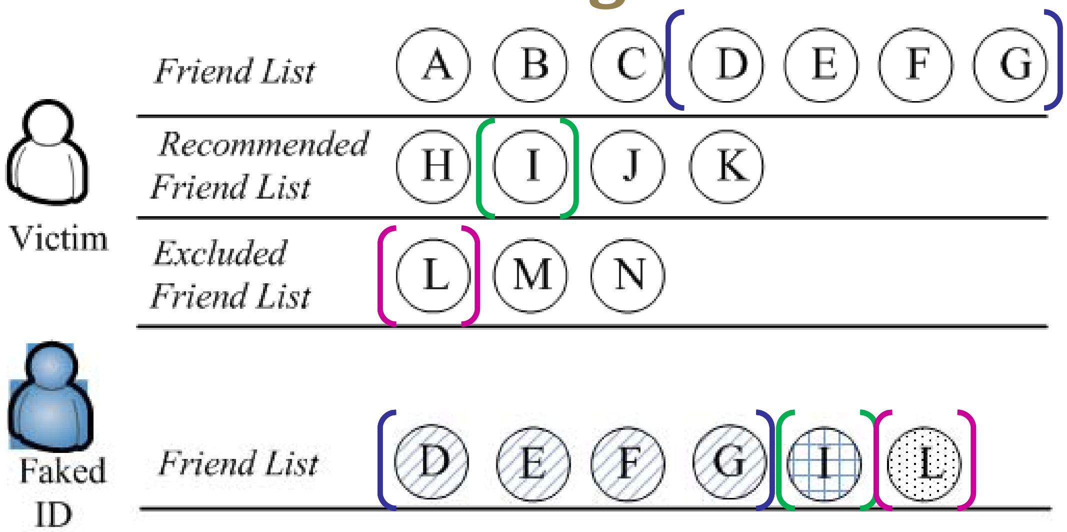
Sub Targets:

1. Attribute Values
2. Privacy Settings

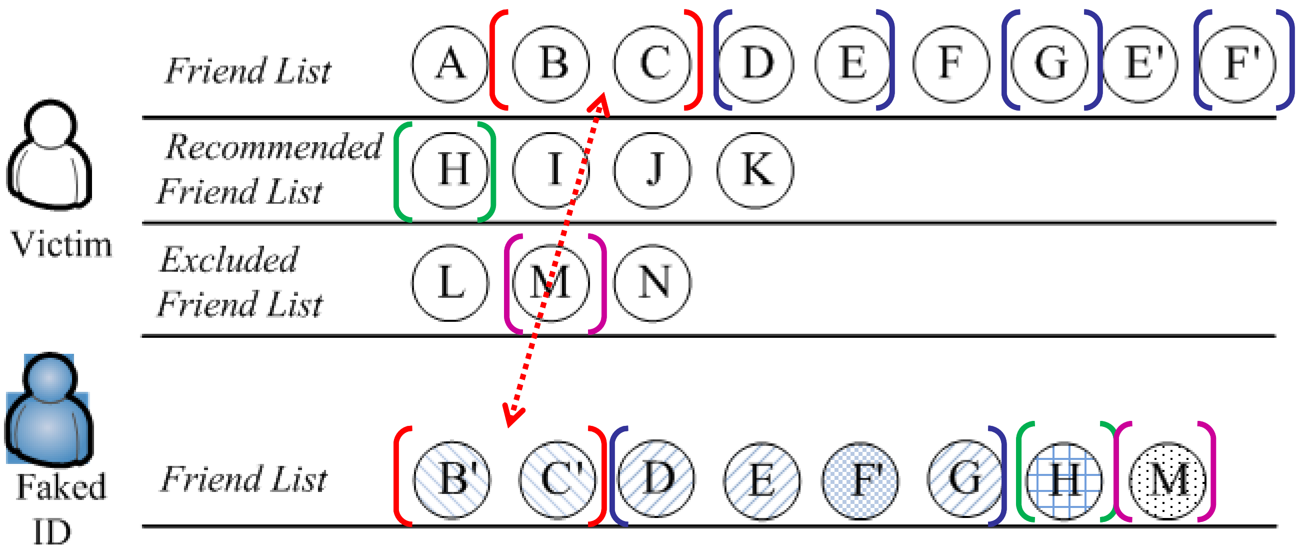


Friend Networks As Target

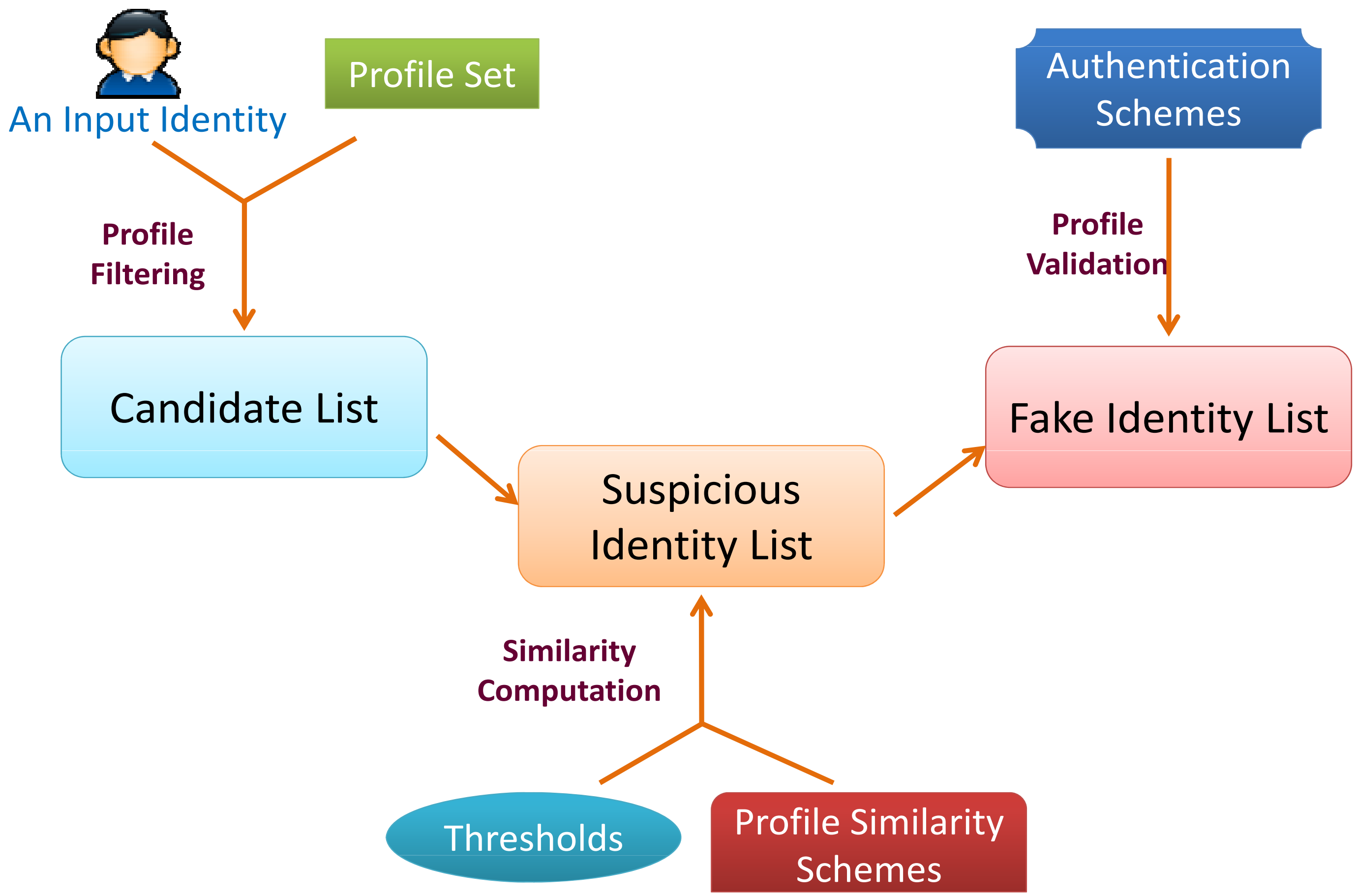
FL
RFL
EFL



FL
RFL
EFL
Faked ID



Cloned Identity Detection



Profile Similarity

Attribute Similarity

$$S_{att}(P_c, P_v) = \frac{SA_{cv}}{\sqrt{|A_c| \times |A_v|}}$$

Basic Principle: Similar Attributes in Two Profiles

Friend Network Similarity

For Basic Profile Similarity (BPS)

$$S_{bfn}(P_c, P_v) = (\alpha S_{ff} + \beta S_{frf} + \gamma S_{fef})$$

Basic Principle:
Mutual Friends in Friend Networks

For Multiple-faked Identities Profile Similarity (MFIPS)

$$S_{mfn}(P_c, P_v) = \alpha(S_{s-ff} + S_{s-cf}) + \beta(S_{s-frf} + S_{s-cfrf}) + \gamma S_{s-fef}$$

Basic Principle:
Similar Friends in Friend Networks

Experiments

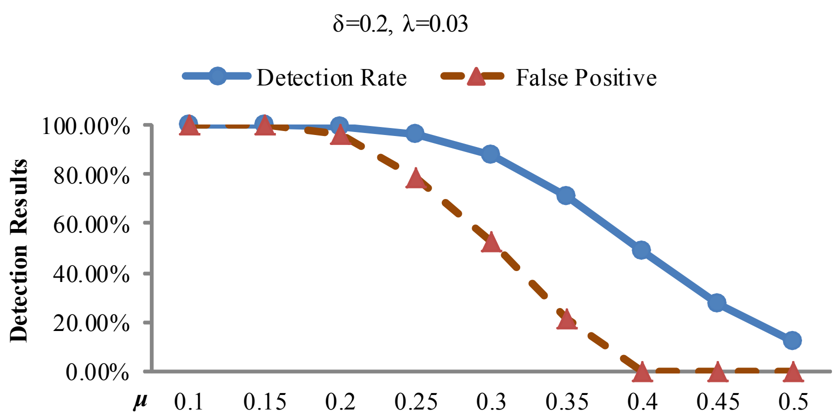


Fig. 3 Relationship between detection results and μ

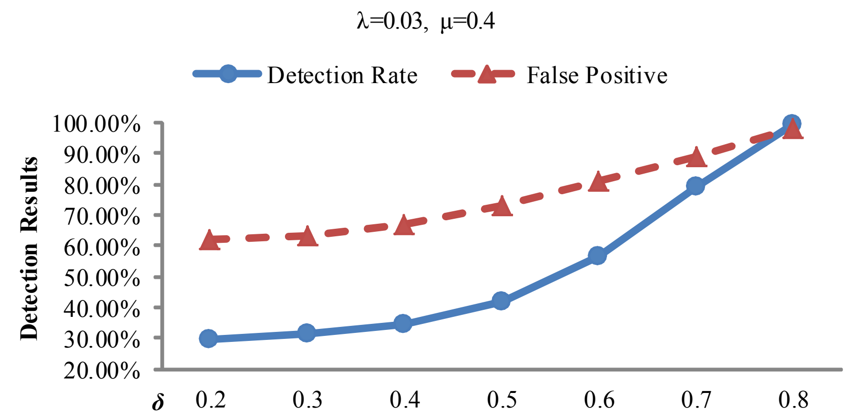


Fig. 4 Relationship between detection results and δ

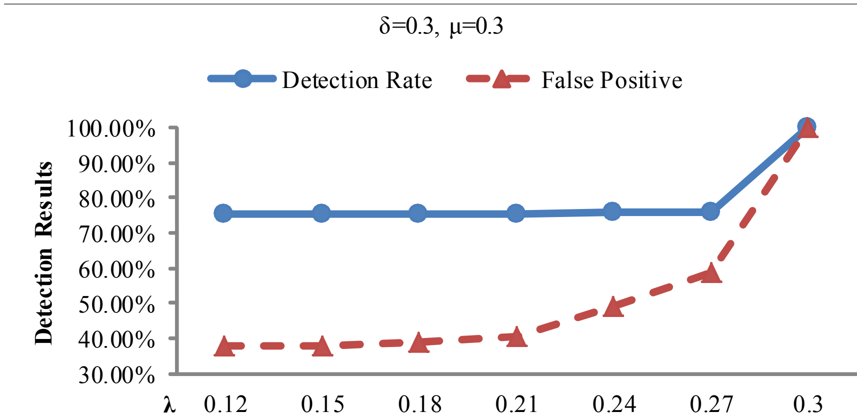


Fig. 5 Relationship between detection results and λ

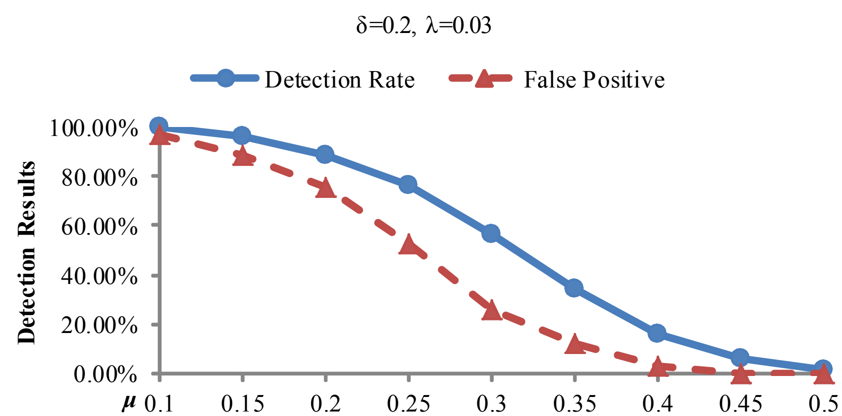


Fig. 6 Friend network similarity on detection results

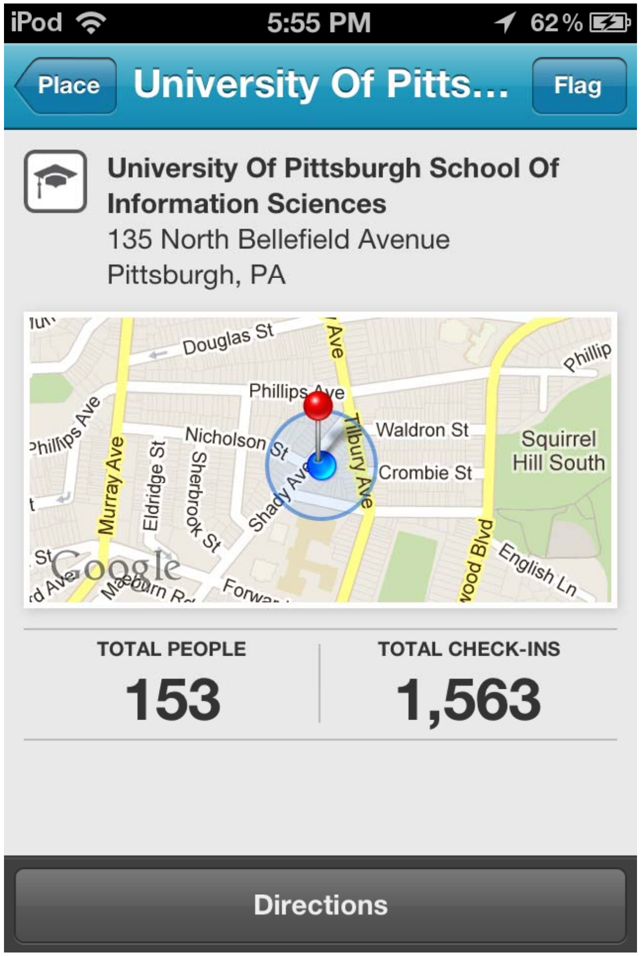
Infer User's Profile Information

- Assumptions: Friends tend to share the same interests
- Inferring a targeted user's private attribute based on his/her friends' public attributes
- Example:
  - A user hides his education and occupation from the public
  - Many of a user's friends are current students at the University of Pittsburgh
  - Inference: University of Pittsburgh, Student

Mislove Alan, Viswanath Bimal, Gummadi Krishna P, Druschel Peter, You are who you know: inferring user profiles in online social networks, WSDM'10

Venue Attacks in LBSNs [2]

- Venue Attributes
  - Creator
  - Owner
  - Name
  - Address
  - Geo-location
  - Category
  - Statistical Information - Owner
  - Promotion/Coupon (Set by Owner)



Malicious Venue Creation Attack

- **ANY** user can create **ANY** type of a venue without being subjected to any **AUTHENTICATION** and the **AUTHORIZATION** from the actual owner
- Venue Not Created in a LBSN
  - Does not exist in the real world: deceive and confuse users, destroy users' trust for LBSNs
  - Exists in the real world but not willing to share; e.g. home, private place
- Venue Already Created in a LBSN
  - Create a similar venue using a similar/alternative name; e.g., School of Information Sciences - iSchool

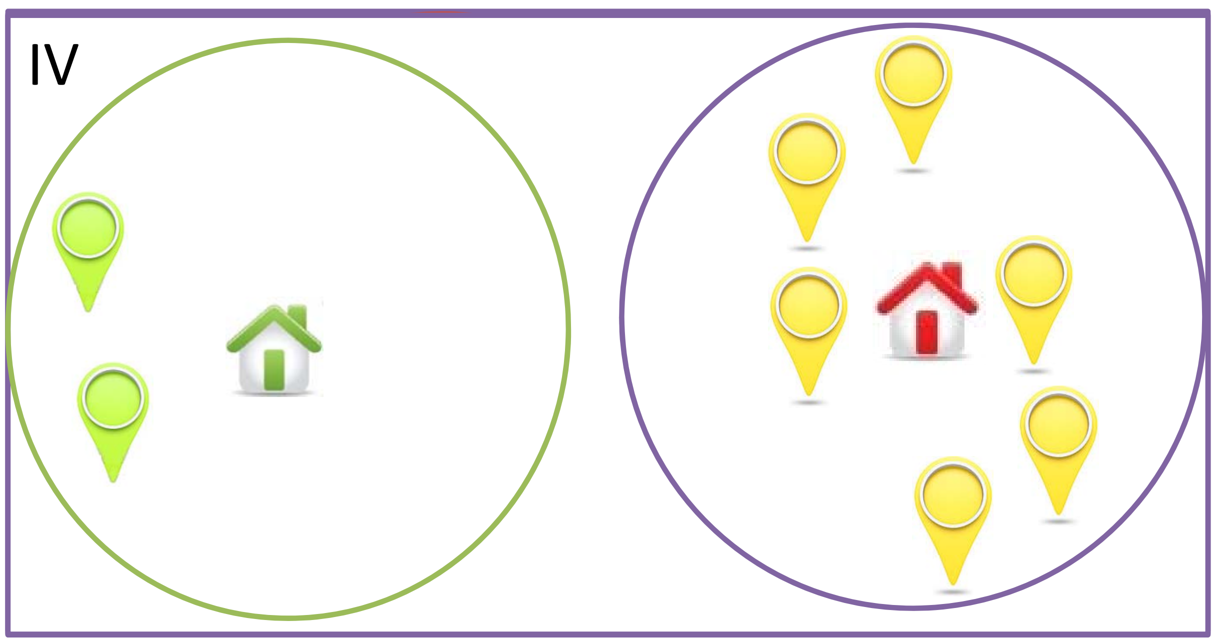
Venue Ownership Hijacking Attack

- Bypass the owner authentication process & become the owner of the created venue
- Owner Authentication in Foursquare, Yelp and Facebook Place
  – Phone number
  – Address
- Impacts
  – Expose customers' visit information: users' privacy
  – Manipulate coupons/promotions: financial loss and/or destroy user trust on the venue
  – Change the address of the venue
  – …

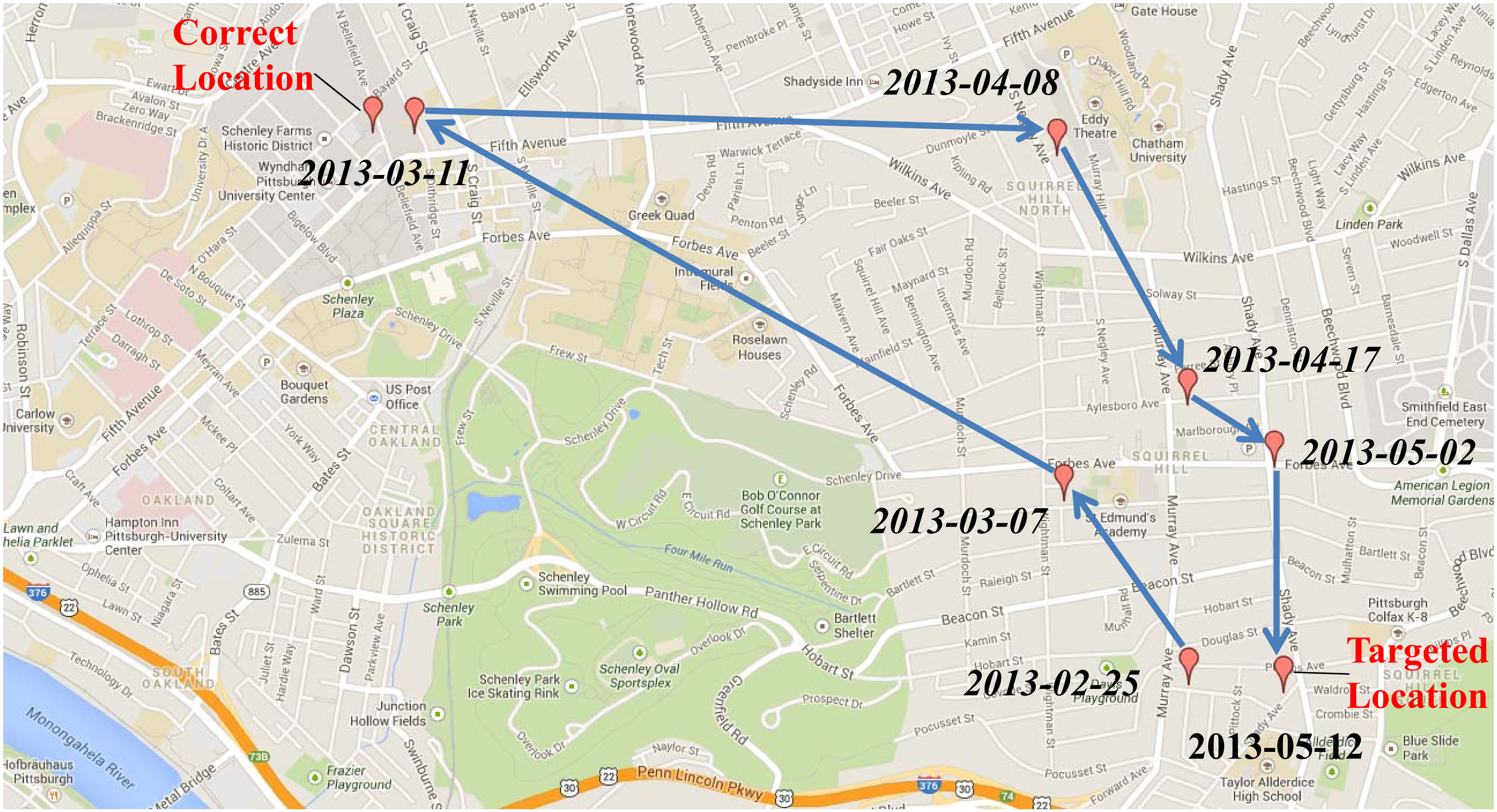
Venue Location Hijacking Attack

- Venue's location is associated with its geo-location not the physical address
- Geo-location is dynamic in terms of possible inaccurate GPS signals
- Location update: the center of all the honest check-ins marked by a LBSN

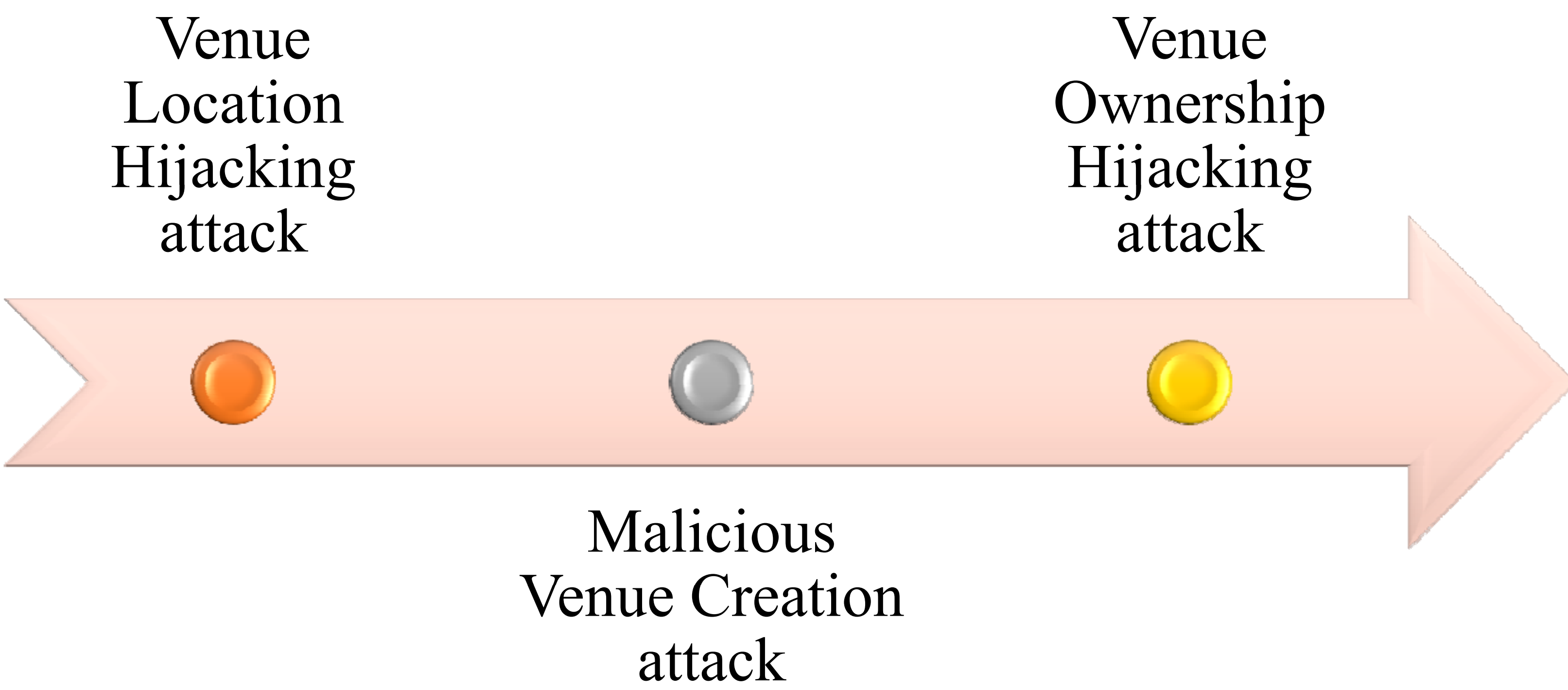
- Users' Honest Check-ins & Marked as Host Check-ins by System
- Users' honest Check-ins & Marked as Dishonest Check-ins by System
- Users' Dishonest Check-ins & Marked as Honest Check-ins by System
- Users' Dishonest Check-ins & Marked as Dishonest Check-ins by System
- Actual Location of the Venue
- Manipulated Location of the Venue

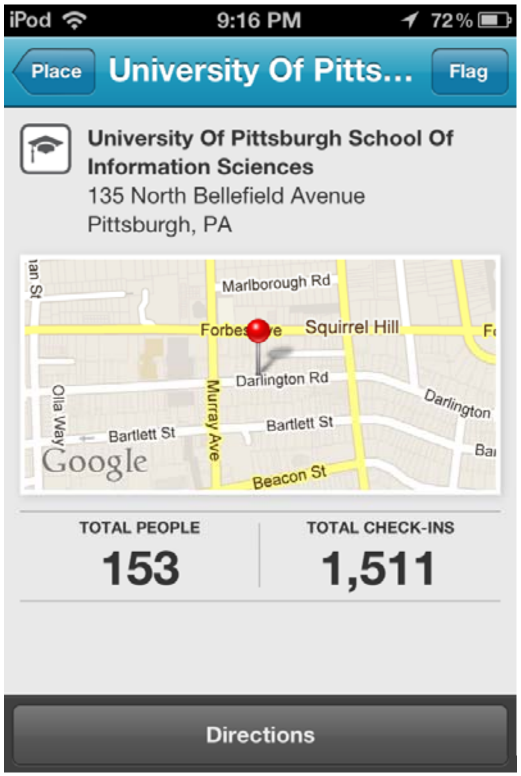


The Movements of the Locations of the LERSAIS Lab

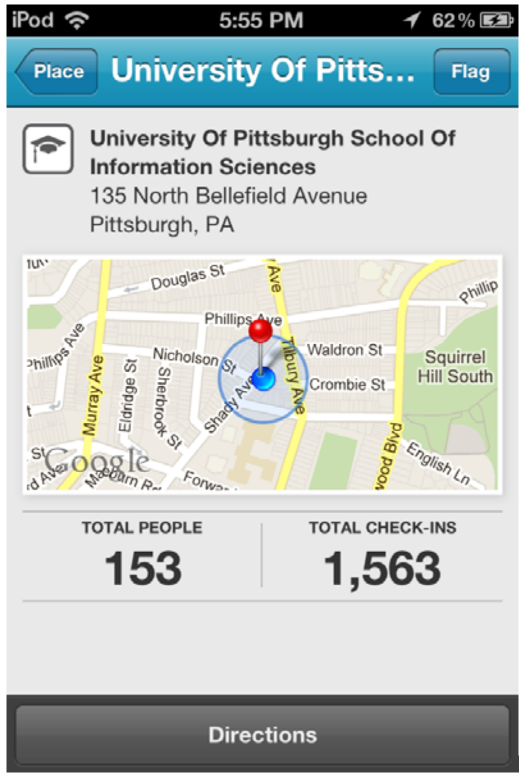


Combined Venue Attacks

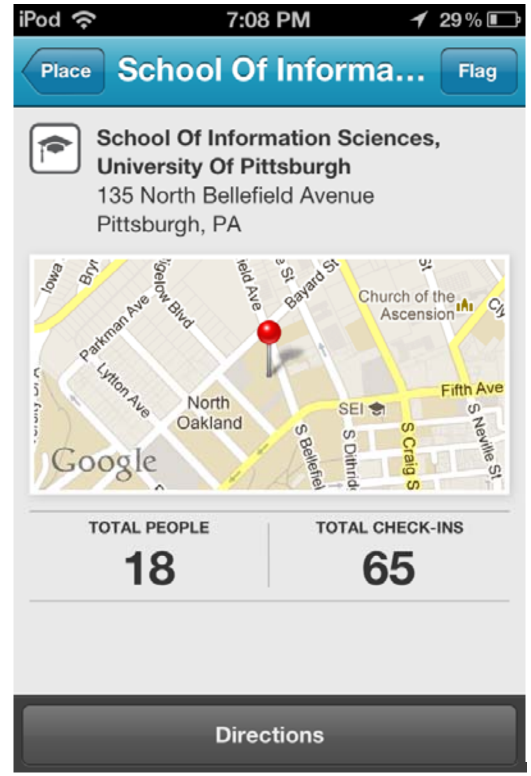




Moved 2 Miles away in May, 2012



Moved 3 Miles away in July, 2012



New Venue Created & Its Check-ins in August, 2012

Investigation of User's Social Network & Posts

User's Social Network
- Friend Lists
- Mutual Friends

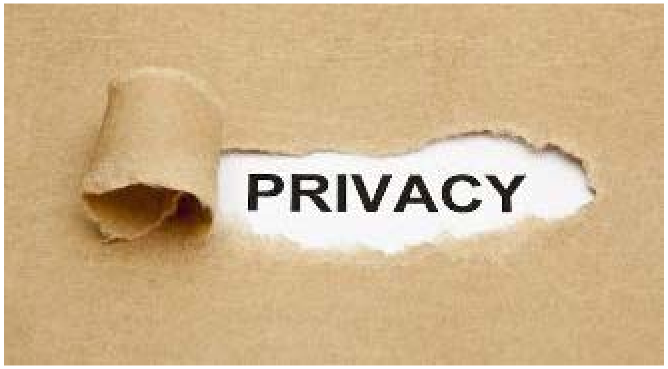
1. Mutual-friend Based Attack
2. Friend Inference Attack

User's Posts
- Photos
- Check-ins

Resource Sharing Issues

Issues Related to Users' Friend Lists

- Importance of the friend list
- What a user's friends reveals
  - Family, Work, Income, Reputation, Religion…
  - Used for Identity Clone Attacks
  - Used for Inferring Private Attributes

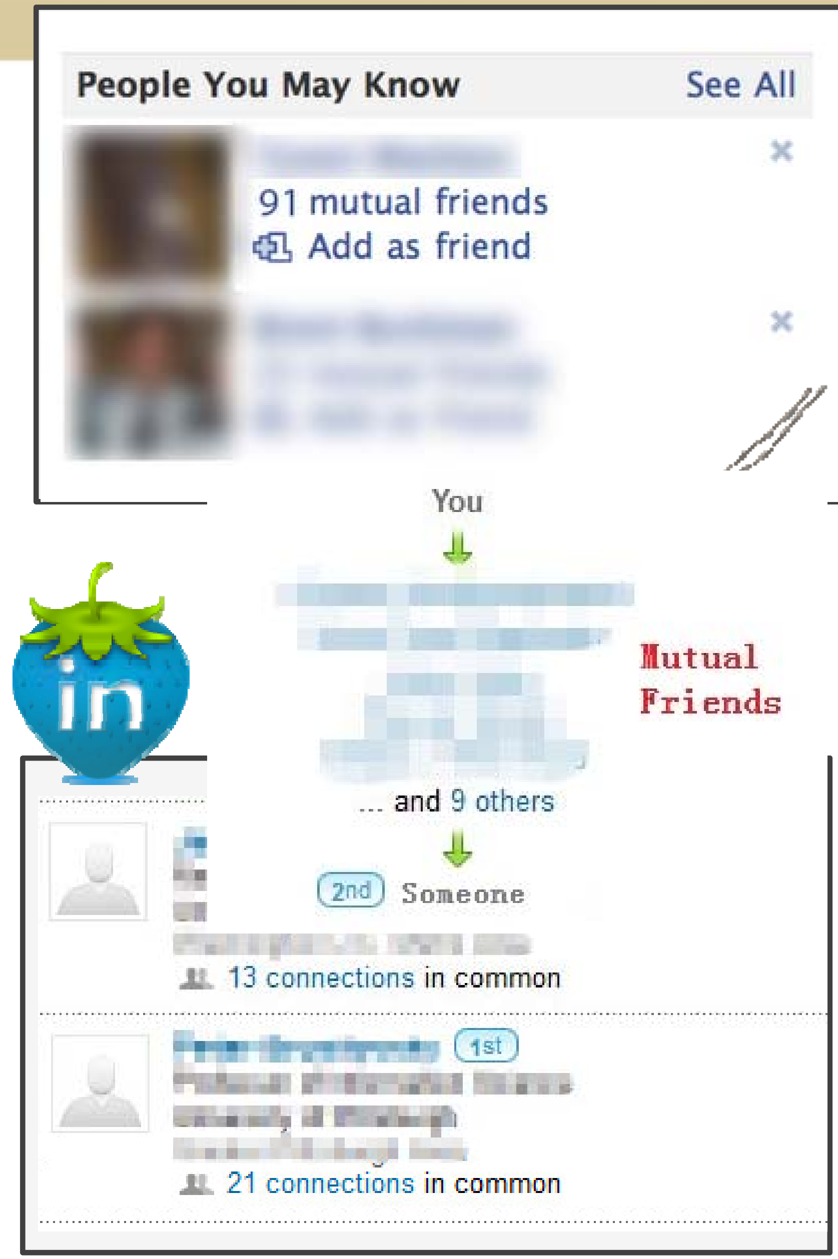


Attacks - Expose a User's Social Network

- Mutual-friend based Attack
- Friendship Identification and Inference Attack

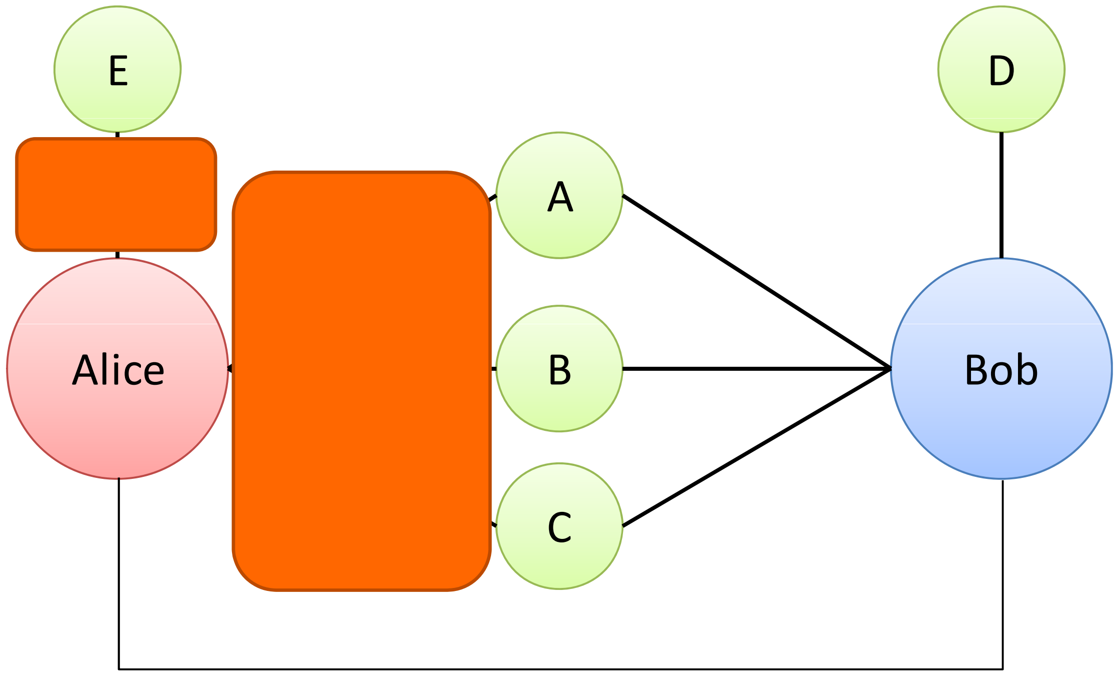
Mutual Friend Feature

- Show mutual friends between two users
- Useful feature, *e.g.* Friend Recommendation, Friend Introduction



Lack of the Access Control Mechanism !

Attack Example



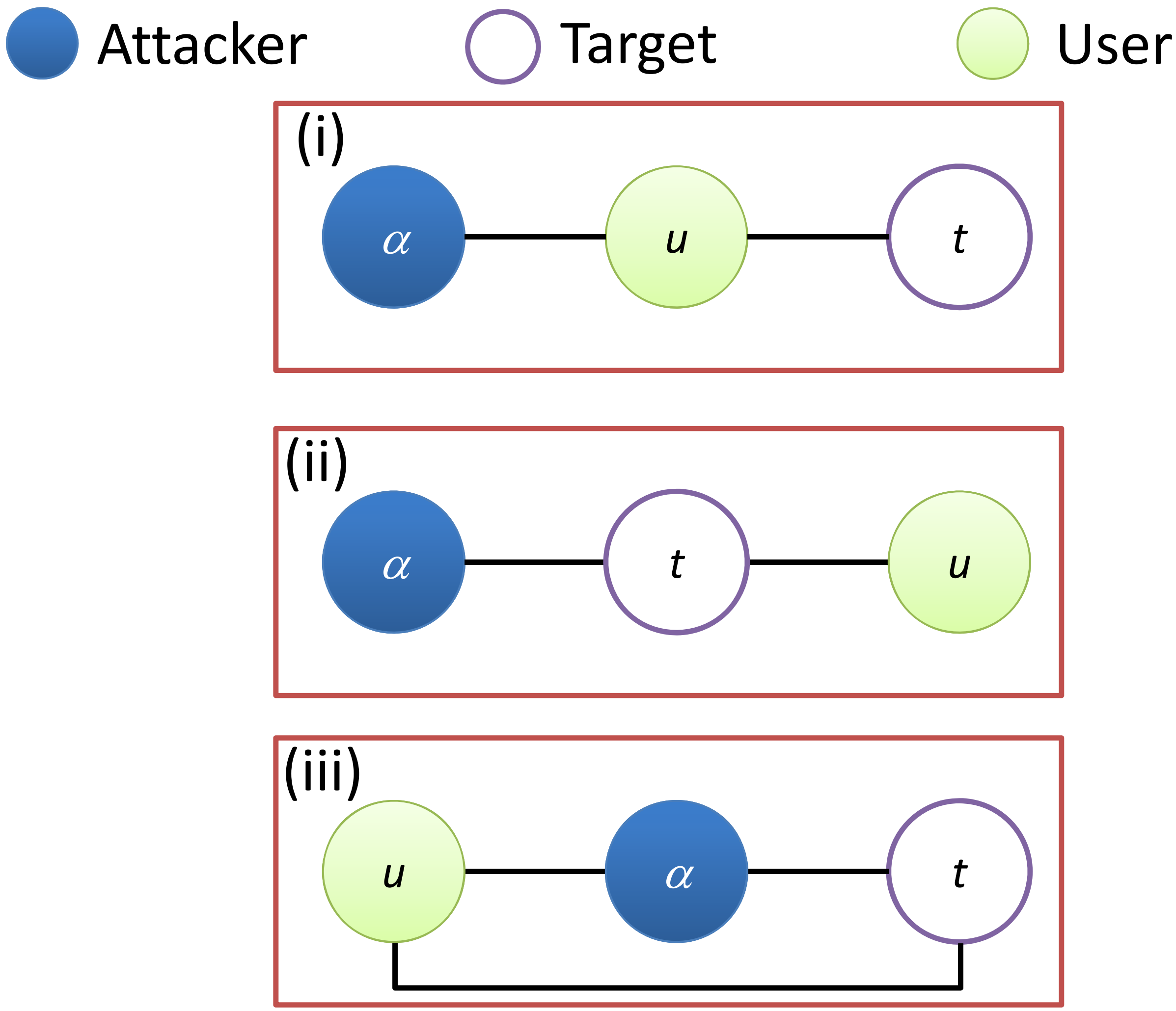
Definition of a Mutual-friend based Attack

- **Attacker (*a*)**
- **Target (*t*):**
  - *t has privacy settings for* ***a***
  - ***a does not*** *know* ***t****'s friends and distant neighbors*
- **Knowledge of an Attacker (Assumptions):**
  - ***a*** *knows his friend list*
  - ***a*** *can find* ***t***
  - *For each* ***u*** *that* ***a*** *can find,* ***a*** *can* ***query*** *MF(****a****,* ***u****)*
- **Mutual-friend based Attack:**
  - *at least one of* ***t****'s friends and/or distant neighbors are* ***exposed*** *to* ***a*** *by querying mutual friends*

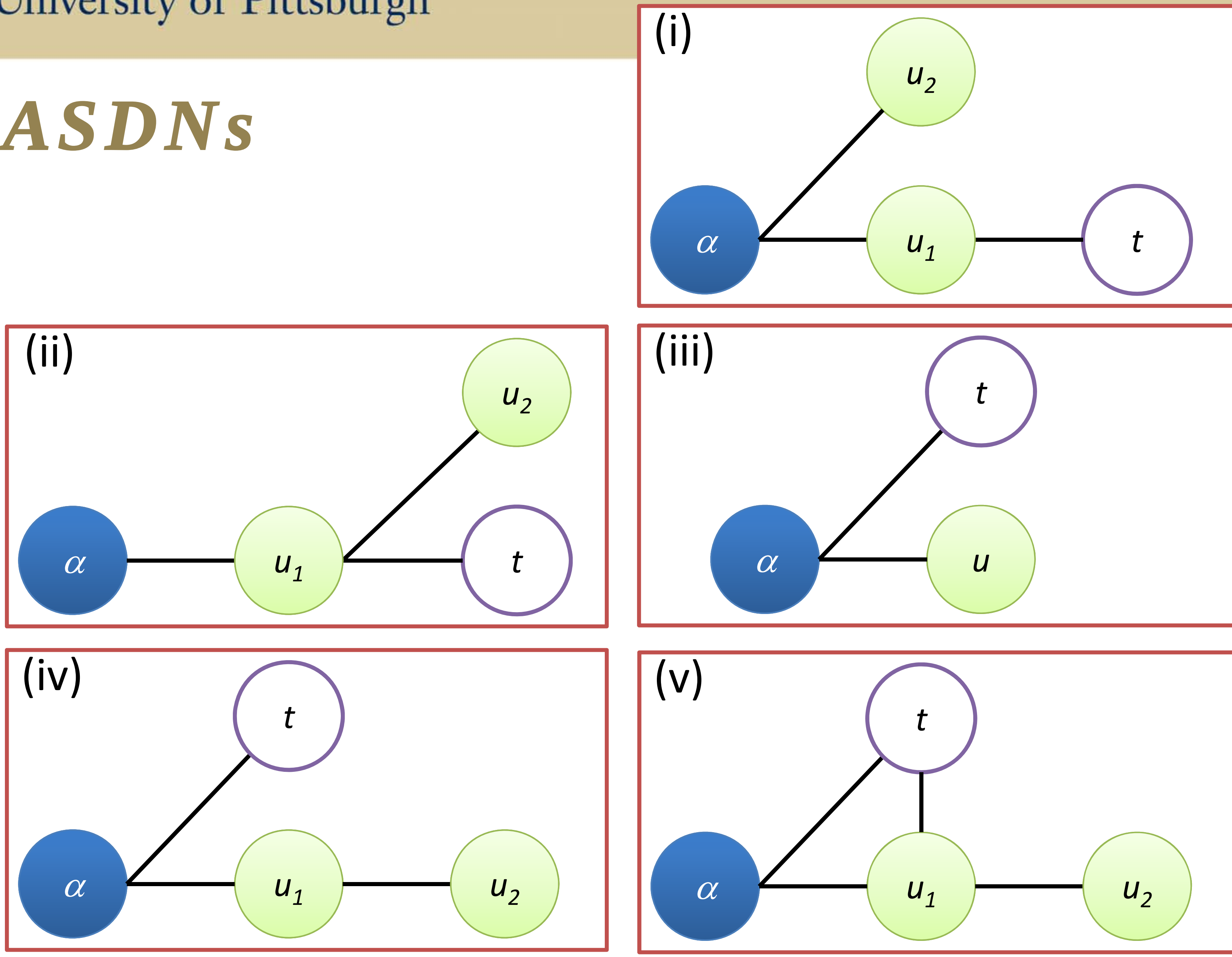
Attack Structures

- Mutual-friend based Attack
  - Conduct various well-designed mutual friend queries
- Queries are designed based on attack structures
  - Types of Attack Structures
    - Used to identify a target's friends (***BASFs***)
    - Used to identify the target's distant neighbors (***BASDNs***)

BASFs



BASDNs


(i)
u_2
α
u_1
t
(ii)
u_2
α
u_1
t
(iii)
t
α
u
(iv)
t
α
u_1
u_2
(v)
t
α
u_1
u_2

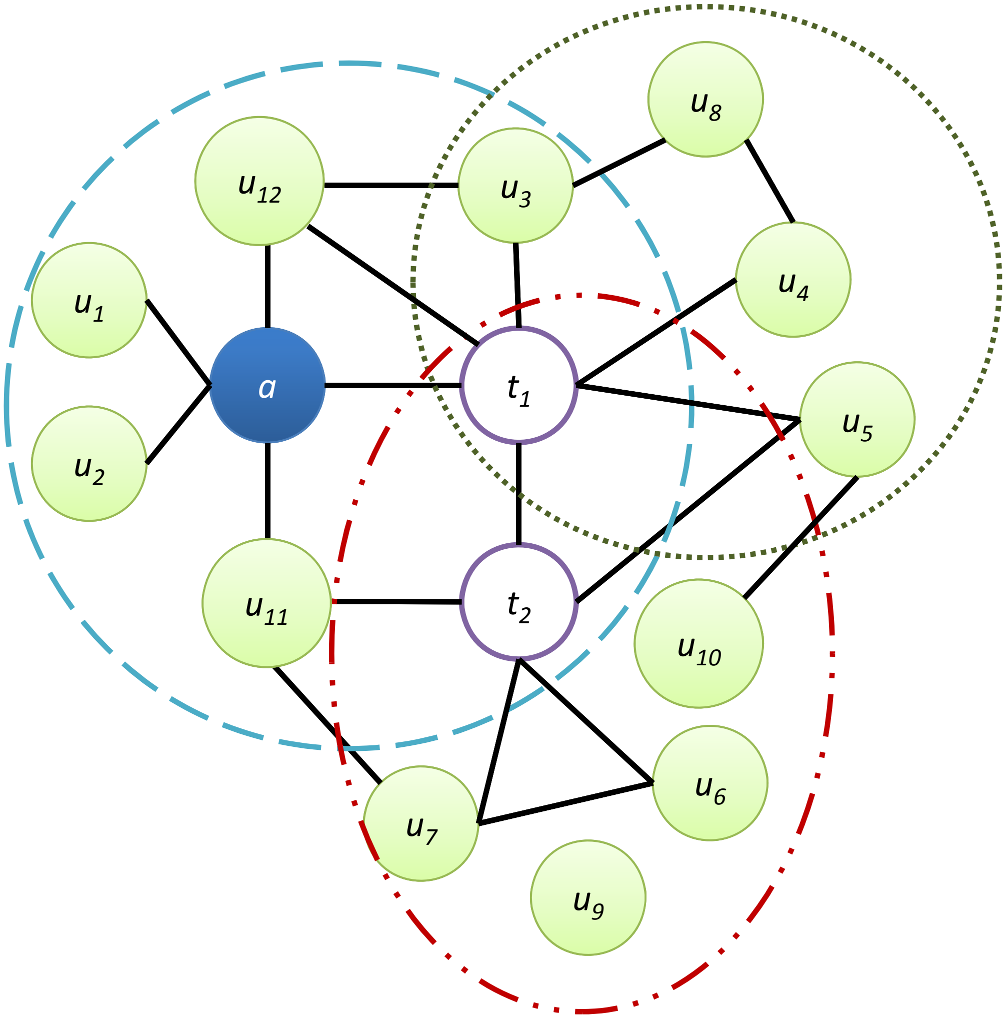
Specific-Target Attack

- Attacker (***a***) has a specific target (***t***)
- Goal:
  - try to find out all the attack structures related to ***t***
  - find out as many ***t***'s friends and distant neighbors as possible
- Attack Steps
  1) Mutual friends between ***a*** & ***t***; ***a***'s friends;
  2) Need a user set *U*
     - a randomized user set (Type 1)
     - a community /group including ***t*** (Type 2)
  3) Find out the attack structures and query mutual friends based on them

Exploration Attack

- Attacker (***a***) has no specific target
- Goal
  - explore users who can be compromised
  - find out only one friend or distant neighbor of a user
- Attack Steps:
  1) Need a target set ***T***
     - a randomized user set (Type 1)
     - a community / group including ***a*** (Type 2)
  2) For each ***t*** in ***T***,
     - if there is one attack structure involving ***a*** & ***t***
     - ***t*** can be compromised

Hybrid Attack



1. Launch the exploration attack using attacker's community

2. Assume the attacker is interested in attacking t_1 and t_2, he chooses them as specific targets

3. Launch specific-target attacks for t_1 and t_2 using t_1's community and a user set about t_2

Specific-target Attacks

	Randomized Group with 100 Users	Target's Group
Average Exposed Friends of a Target	2.6 (10.2%)	12.4 (48.9%)
Average Exposed Distant Neighbors of a Target	24.7	42.0

Exploration Attacks

	Randomized Group with 100 Users	Attacker's Group
Average Compromised Users	12.3 (12.3%)	52.65 (61.0%)

Hybrid Attacks

- *Exploration + Specific-target*
  - ❖ Launch Exploration attacks using the attacker's groups
  - ❖ Choose two specific targets who have the most mutual friends with the attacker from the Exploration attack results
  - ❖ Launch Specific-target attacks for selected targets using targets' groups

	Results of Hybrid Attacks	Results of Specific-target Attack using the target's group
Average Exposed Friends of a Target	19.4 (73.2%)	~ 12 (49%)
Average Exposed Distant Neighbors of a Target	48.3	~ 42

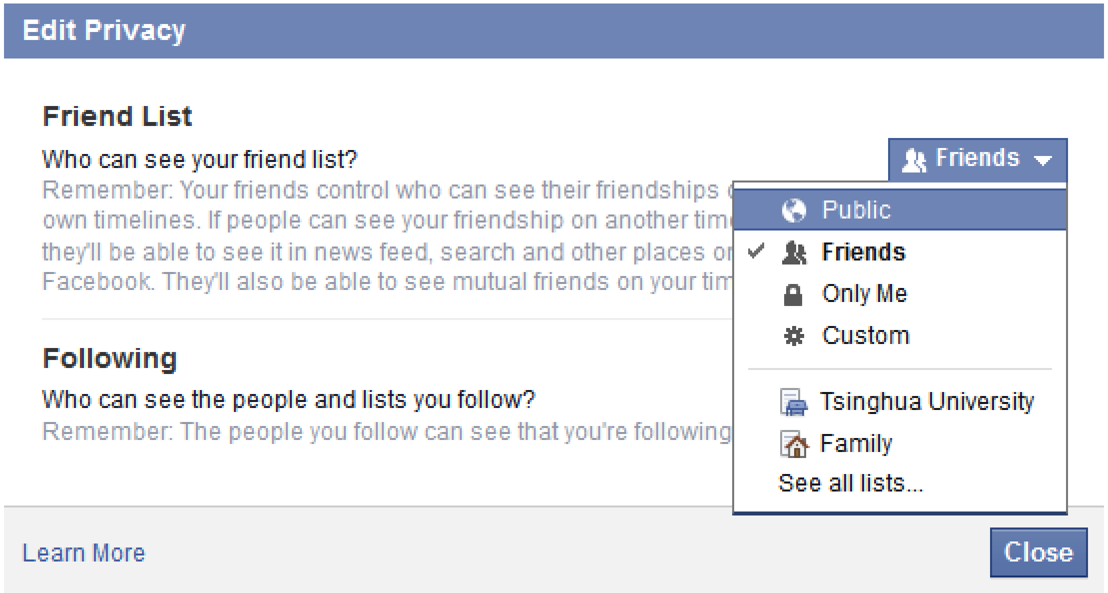
Defense Approaches

- Reason
  - ❖ no restriction for querying mutual friends
- Defense approaches
  - ❖ Hide user profile
  - ❖ Access control to query mutual friends

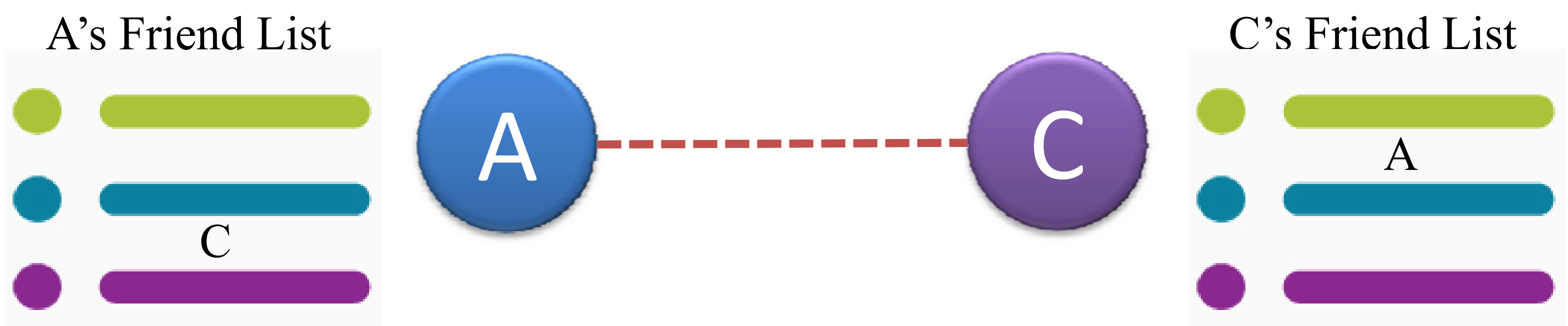
Friendship Identification & Inference Attack

- Users' Privacy Settings for Friend Lists
  - Private
  - Friends w/o an excluding list
  - Public

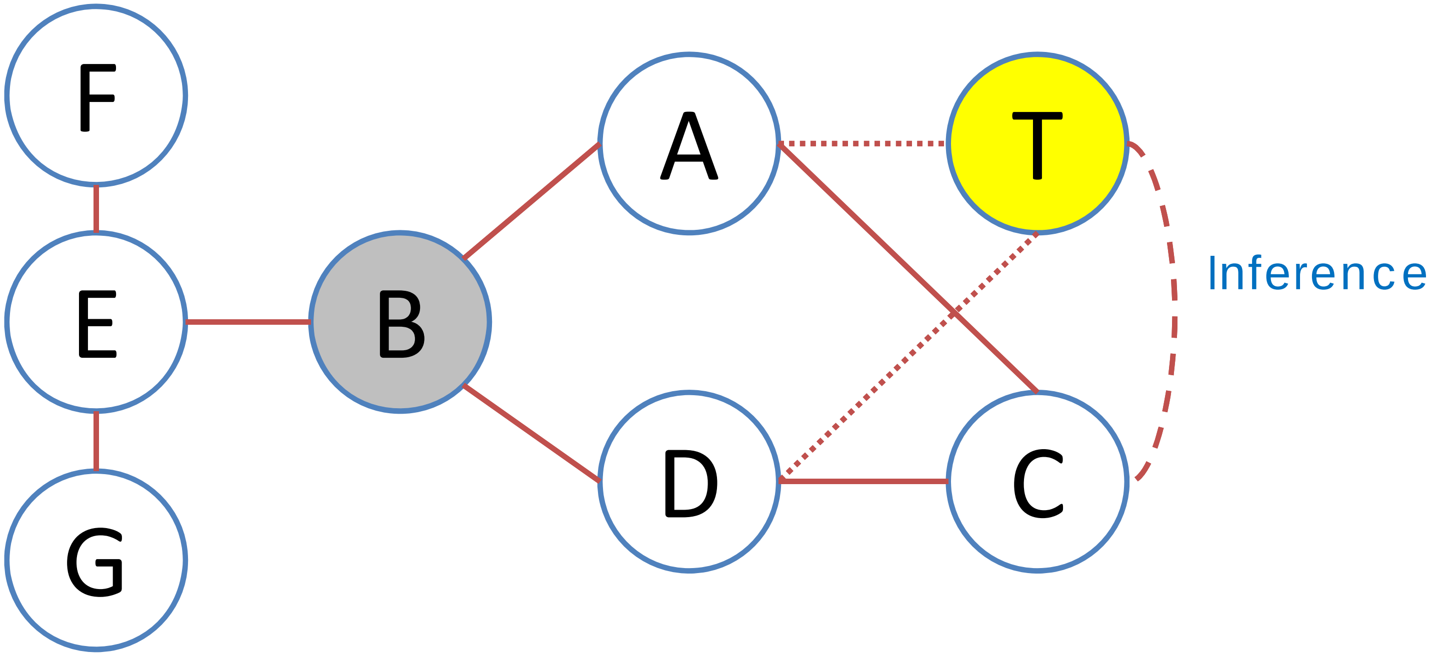
Consistent Among Users?



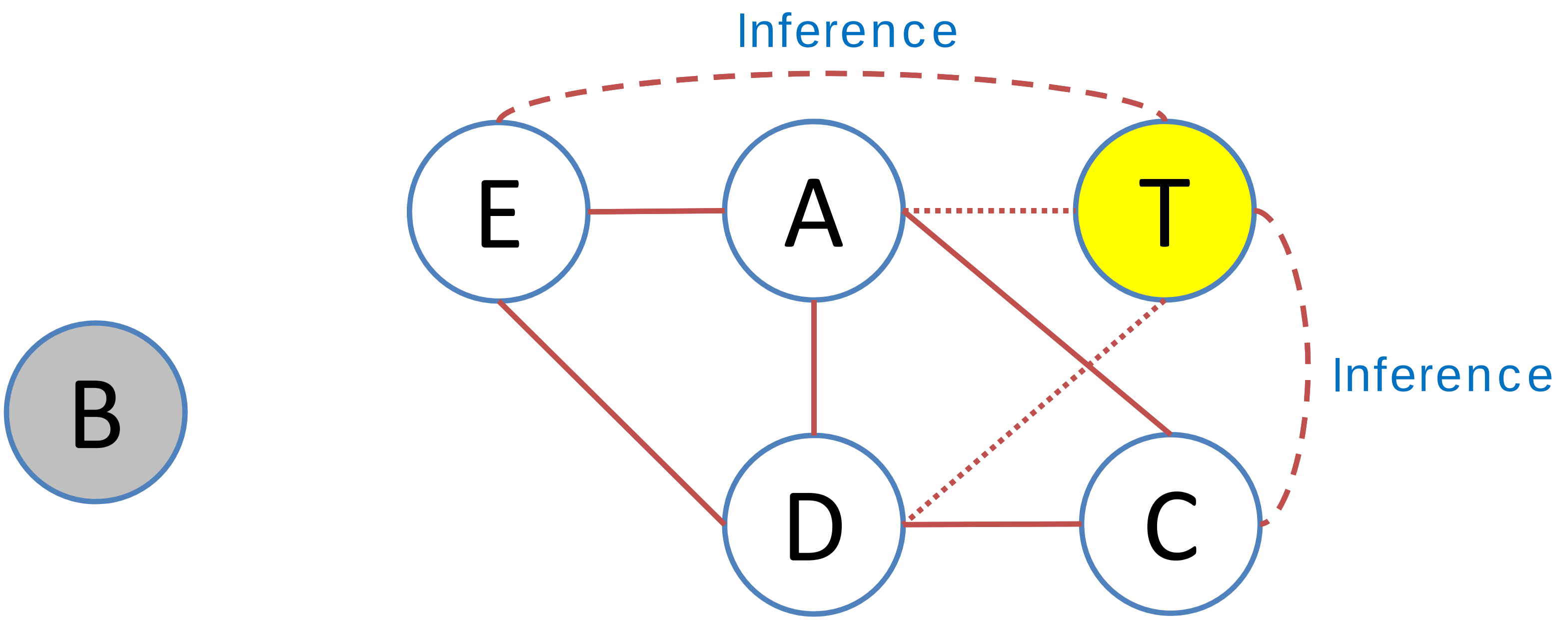
Inconsistent Policies



Inconsistent Preferences Example -1



Inconsistent Preferences Example -2



Attack Definition

A social graph $G(V, E)$; An adversary $\boldsymbol{b} \in V$; A target $\boldsymbol{t} \in V$

Assumption on Target t's Privacy Setting:

t defines a policy that does not authorize ***b*** to see ***F*(*t*)**

Assumptions on Adversary's Initial Knowledge:

- ***b***'s initial attack knowledge $\boldsymbol{K(b) = (V_{kb}, E_{kb})}$ is constructed based on friend lists visible to him
- ***t*** is included in $\boldsymbol{V_{kb}}$

Privacy Attack:

t is a victim of an *FII* attack launched by ***b*** if ***b*** can identify and correctly infer at least ***e*** friends of *t*'s friends based on the proposed **random walk based link predictions** on ***K*(*b*)**.

No Additional Action

Attack Schemes

- **One attacker node & one target**
  - **Scheme 1**: Adversary chooses a number of users, who are the most likely to be friends of a target, at one time
  - **Scheme 2**: Adversary choose only one user, who is the most likely to be friend of a target at one time; add such a friendship link to the network and launch the attack again…
  - **Scheme 3**: Adversary first attacks other users, who are close to the target; add identified and inferred friendship links to the network; then attack the target

Attack Schemes (cont.)

- **Multiple attacker nodes & one target**
  - ❖ Combine the attack knowledge (segments of the network) from different attacker nodes to be a more completed segment of the network
- **Topology of the entire social network (multiple attacker nodes & multiple targets)**
  - ❖ Attack the most vulnerable targets first

Results of attack scheme 1 in the three datasets

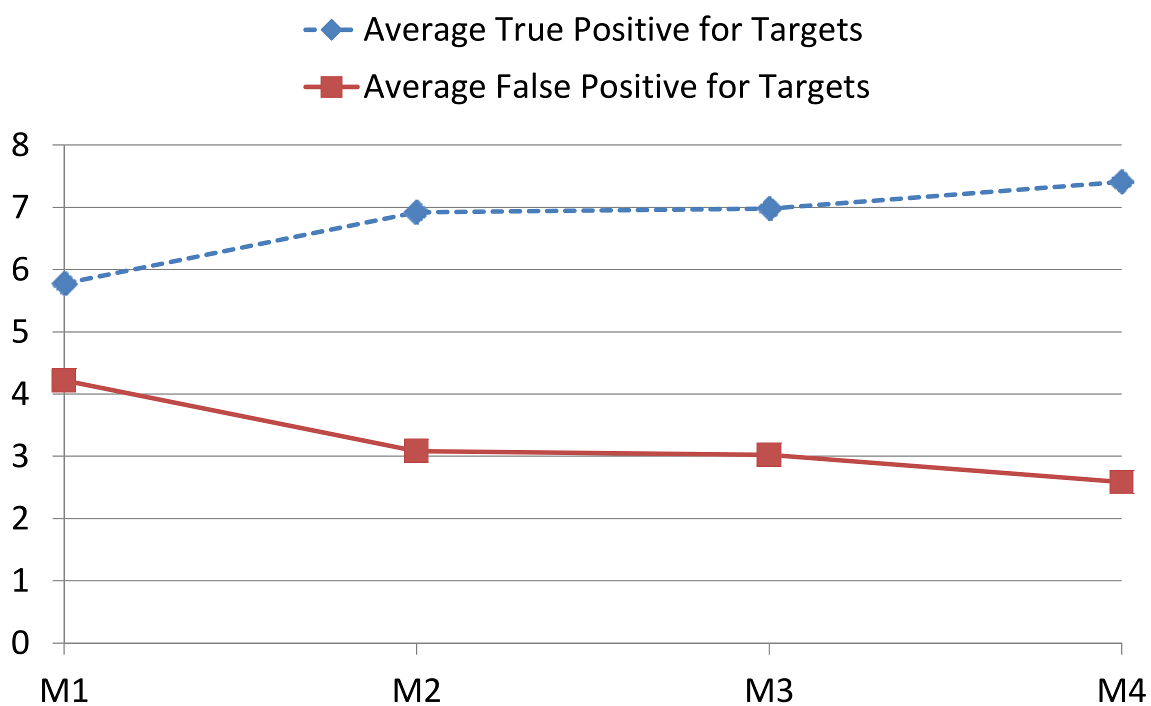
	Relationship Between an Attacker Node & a Target	D1	D2	D3
Average True Positive for Attacker Nodes (Average RC)	Friend	7.82 (78.2%)	4.71 (47.1%)	5.48 (54.8%)
	2-distant Neighbor	5.78 (57.8%)	2.85 (28.5%)	3.25 (32.5%)
	More than 2-distant Neighbor	4.03 (40.3%)	3.13 (31.3%)	3.19 (31.9%)
Average False Positive for Attacker Nodes (Average RIC)	Friend	2.18 (21.8%)	5.29 (52.9%)	4.52 (45.2%)
	2-distant Neighbor	4.22 (42.2%)	7.15 (71.5%)	6.75 (67.5%)
	More than 2-distant Neighbor	5.97 (59.7%)	6.87 (68.7%)	6.81 (68.1%)

Results of attack schemes in the three datasets

	Attack Scheme	D1	D2	D3
Average True Positive for Attacker Nodes	Attack Scheme 1	5.88	3.56	3.97
	Attack Scheme 2	5.92	3.78	4.22
	Attack Scheme 3	6.91	4.85	5.32

Attack results using multiple attacker nodes in D3

- M1: **one** friend, **one** 2-distant neighbor, **one** more than 2-distant neighbor
- M2: **three** friend, **one** 2-distant neighbor, **one** more than 2-distant neighbor
- M3: **three** friend, **five** 2-distant neighbor, **five** more than 2-distant neighbor
- M4: **five** friend, **five** 2-distant neighbor, **five** more than 2-distant neighbor

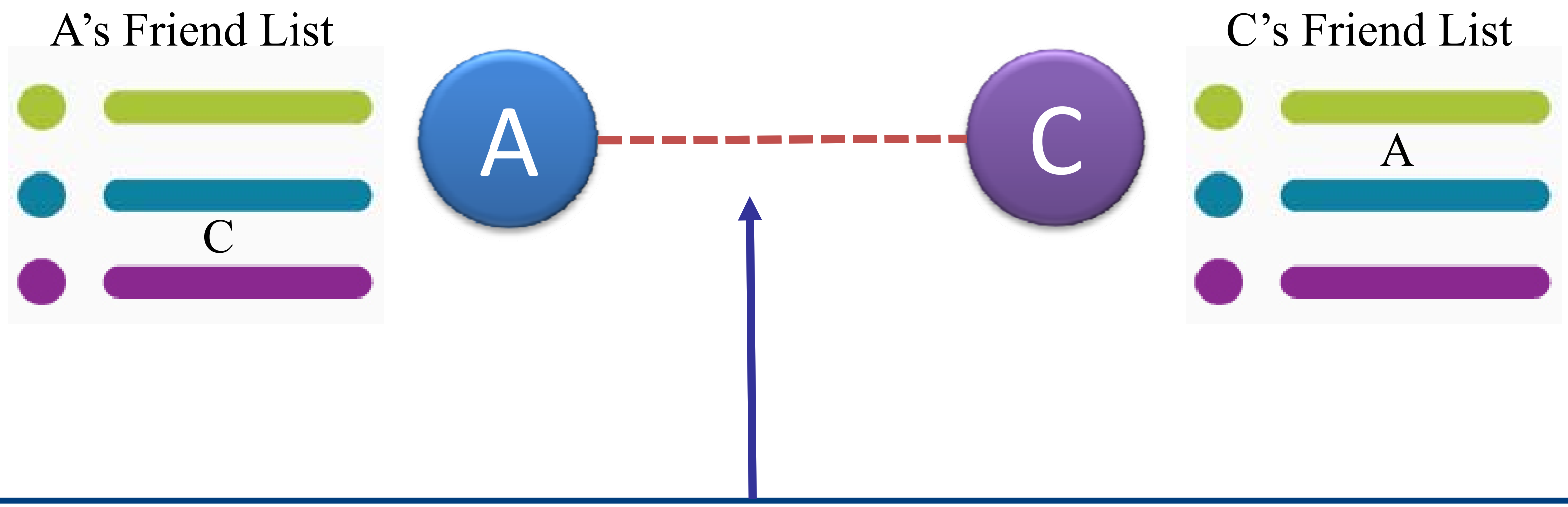


Results from *FII* attacks on the topology of the entire network in D3

Each attack is repeated in 10 times

	After Launching 1K FII Attacks	After Launching 5K FII Attacks	After Launching 10K FII Attacks	After Launching 20K FII Attacks
Average Correctly Inferred Friendship Links (Percentage of All Friendship Links)	4113.4 (7.76%)	17948.3 (33.88%)	32231.4 (60.84%)	37891.2 (71.53%)
Average Incorrectly Friendship Links	4528.6	13314.4	21653.2	25472.3

Defense Approaches



- Squicciarini et al. -> voting algorithm & game theory
- Hu et al. -> Label Privacy Level, minimize privacy risk & sharing loss

Issues Related to Users' Posts

- Photos
  - A photo includes multiple individuals
  - One of them posts it in his/her wall
  - Other may be upset
- Check-ins (LBSNs) [2]
  - A user exposes where and when he is
  - A user exposes where his lives
  - A user's friend or other people expose the user's location related information
- Existing Access Control mechanisms cannot address all of these problems [5]

Other Issues – Email Address as Identity [7]

- Too many online systems adopt a user's email address as the user's identity
- Caused and causing many threats
  - Email address is not considered to be a private information
  - Easy to guess a user's identity in a online system
  - More vulnerable for online password cracking
    - Share the same passwords
    - Avoid the limits of fail login times
  - Cracking one email address = Cracking related online accounts associated with this email address

Social Media Landscape *2013*

References

1) Lei Jin, Hassan Takabi, Xuelian Long, James B.D. Joshi, Exploiting Users' Inconsistent Preferences in a Social Network System to Discover Private Friendship Links, WPES'14.
2) Lei Jin and Hassan Takabi, Venue Attacks in Location-Based Social Networks, GeoPrivacy'14.
3) Lei Jin, James B.D. Joshi, Mohd Anwar, Mutual-friend Based Attacks in Social Network Systems, Computer & Security, v. 37, pp-15-30, 2013.
4) Lei Jin, Xuelian Long, James B.D. Joshi, Towards Understanding Residential Privacy by Analyzing User' Activities in Foursquare, BADGERS'12.
5) Lei Jin, Xuelian Long, James B.D. Joshi and Mohd Anwar, Analysis of Access Control Mechanisms for Users' Check-ins in Location-Based Social Network Systems,WICSOC 2012.
6) Lei Jin, Hassan Takabi and James B.D. Joshi, Towards Active Detection of Identity Clone Attacks on Online Social Networks, CODASPY 2011.
7) Lei Jin, Hassan Takabi and James B.D. Joshi, Security and Privacy Risks of Using E-mail Address as an Identity, PASSAT 2010.

Questions?

Thank You!