

# Privacy Issues LinkedIn

CPSC 457/557 - Sensitive Information

October 7th, 2013

Haotian Xu

# Outline

- LinkedIn - statistics and status quo
- Link Reconstruction Attack
- Find hidden connections on LinkedIn
  - Algorithm and my experiment
- Cases
  - 2012 - LinkedIn iOS app collects private info
  - 2013 - LinkedIn “Intro” iOS mail app extension
- Thoughts & Conclusion

# LinkedIn: Statistics & Status Quo [1][2]

- **Statistics**

- Members from 4500 to 259M in 10 years
- 39% users pay for the premium service
- 35% users access the site daily
- 42% users update their pages regularly
- 2.7 million of business pages
- 81% users belong to at least one group

- **Implications**

- Big impact, 1% means two and half million users
- Business information and more attractive to hackers
- Higher privacy expectations from premium users

# Link Reconstruction Attack<sup>[3]</sup>

- One's personal information can be inferred from one's links
- Apply supervised learning algorithms to solve link prediction problem
  - Introduced to SNS by Liben and Kleinberg in 2003
- Michael Fire's paper: based on a classifier trained on a small set of easy-to-compute topological features
  - Quantitative analysis on nodes and connections
  - Feature extraction

# Link Reconstruction Attack

Formally

- $G = \langle V, E \rangle$ , set of edges and vertices
- Edge:  $e = (u, v) \in E$  where  $u, v \in V$
- Training Set:  $\langle E \rangle_{\text{training}}$  - Positive & Negative
- Classifiers: compute the *likelihood* of  $(u, v) \in E$  or  $(u, v) \notin E$  for every two nodes  $u, v \in V$

# Link Reconstruction Attack

## Feature Extraction

- Vertex degree
  - $\Gamma(v) := \{u | (u,v) \in E \text{ or } (v,u) \in E\}$ ,  $\text{degree}(v) := |\Gamma(v)|$
- Common Friends
  - $\text{common-friends}(u, v) := |\Gamma(v) \cap \Gamma(u)|$ , for  $u, v \in V$
- Total Friends
  - $\text{total-friends}(u, v) := |\Gamma(u) \cup \Gamma(v)|$ , for  $u, v \in V$
- Jaccard's coefficient
  - $\text{Jaccard's-coefficient}(u, v) := |\Gamma(u) \cap \Gamma(v)| / |\Gamma(u) \cup \Gamma(v)|$
- Preferential-attachment-score
  - $\text{preferential-attachment-score}(u, v) := |\Gamma(u)| \cdot |\Gamma(v)|$

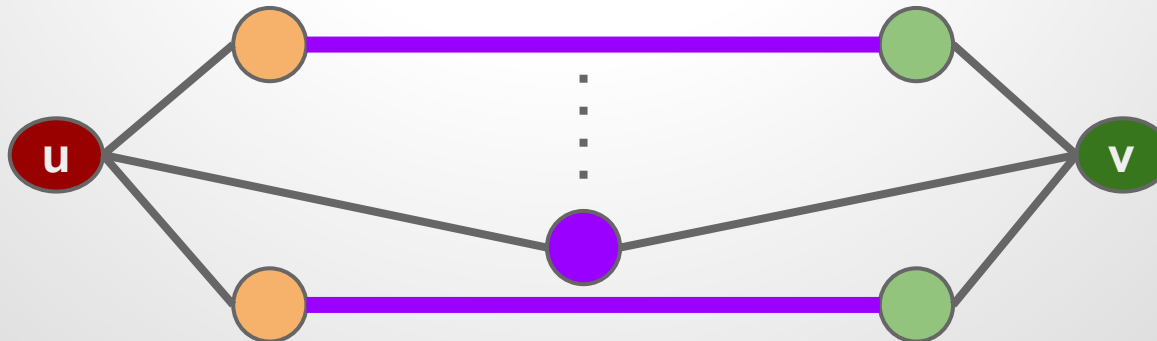
# Link Reconstruction Attack

- Shortest path
  - shortest-path( $u, v$ ) - Dijkstra's, Heuristic  $A^*$ , etc
- Friends Measure

$$\text{friends - measure}(u, v) := \sum_{x \in \Gamma(u)} \sum_{y \in \Gamma(v)} \delta(x, y)$$

Where  $\delta(x, y)$  is defined as:

$$\delta(x, y) := \begin{cases} 1 & \text{if } x = y \text{ or } (x, y) \in E \text{ or } (y, x) \in E \\ 0 & \text{otherwise} \end{cases}$$



# Link Reconstruction Attack

- Experiment in 2012
  - FB - 410 nodes, 635 links, data from web crawler
  - Result has very high *Truth Positive* rate > 98%
  - LinkedIn is comparable to Facebook (if not better)
    - Undirected Connection
    - More cluster distribution on company
    - More features extraction criterions - current and previous companies, groups, employment period

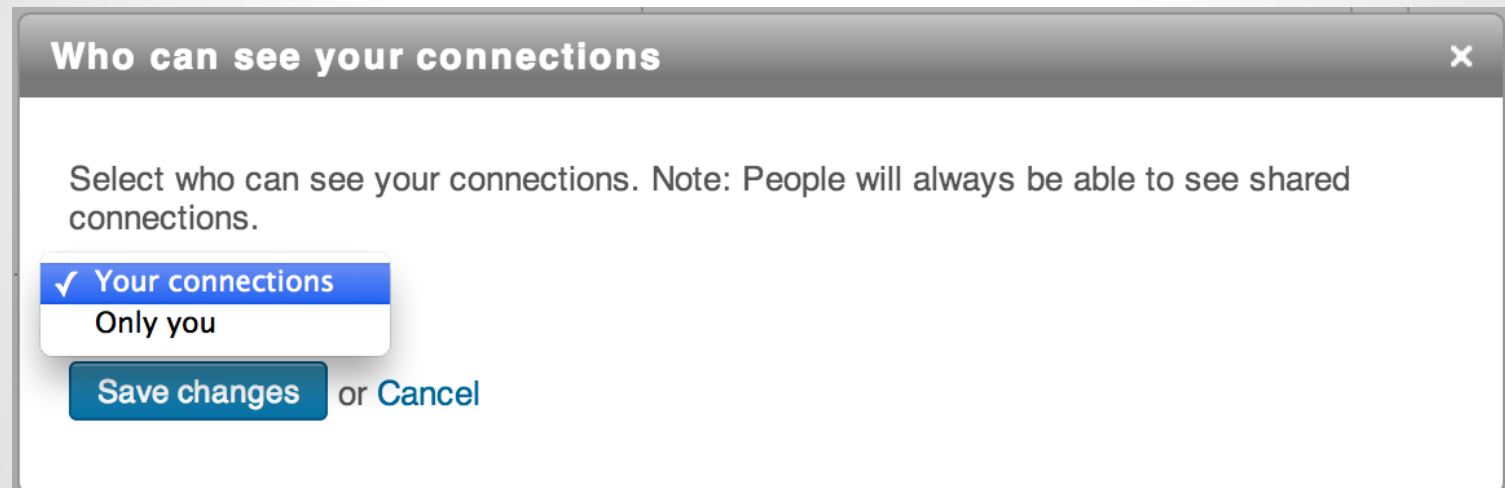
**Table 1.3** Hard Training Set - Classifiers' Highest Results

Dataset	Classifier	Train set size	TP	F-Measure	AUC
Facebook	Rotation Forest	1,270	0.9835	0.9686	0.9981
Academia.edu	Rotation Forest	1,445	0.9127	0.9213	0.9756
Friends & Family	Rotation Forest	563	0.9537	0.9471	0.9831
Student's Cooperation	Rotation Forest	623	0.999	0.9883	0.9998



# Find Hidden LinkedIn Connections<sup>[4]</sup>

LinkedIn's privacy setting



However....

# Find Hidden LinkedIn Connections

- Jessica Staddon - researcher from PARC
- Algorithm:
  1. Attacker **A** establishes connection w/ Target **T**
  2. From **T**'s profile extract attributes
  3. Put those attributes into LinkedIn's search tool
  4. For every returned profile **V**, add into set **Ct** if **T** is a shared connection between **A** and **V**
  5. Return set **Ct**
- Problem
  - LinkedIn limits query results to 100
  - Attacker's own network

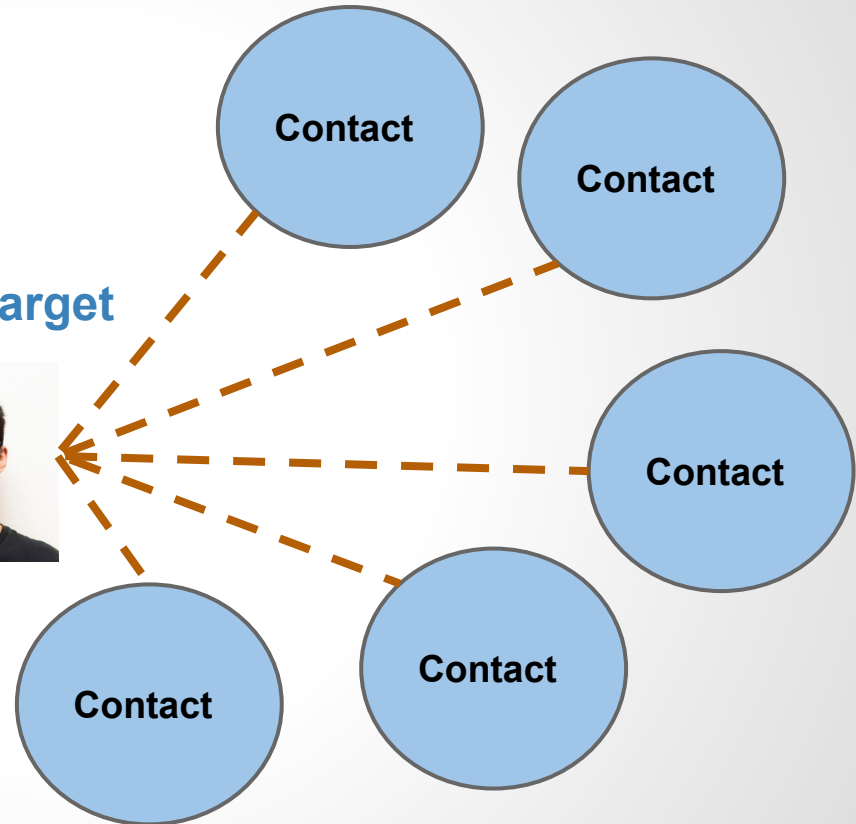
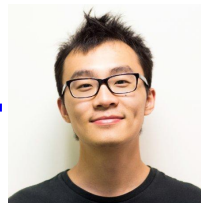
# Find Hidden LinkedIn Connections

Experiment:

Ding Yi - Attacker



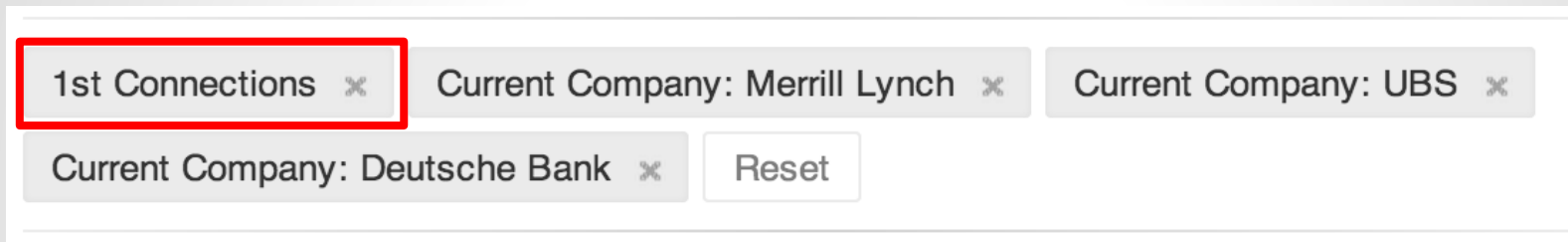
Haotian - Target



Goal: Find Haotian's old colleagues (hidden)

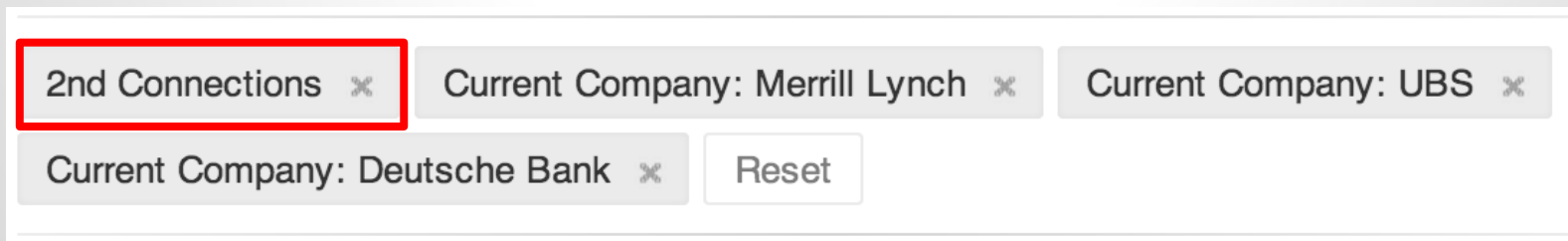
# Find Hidden LinkedIn Connections

Target's account (Haotian) search criterias:



1st Connections x Current Company: Merrill Lynch x Current Company: UBS x  
Current Company: Deutsche Bank x Reset

Attacker's account (Ding Yi) search criterias:



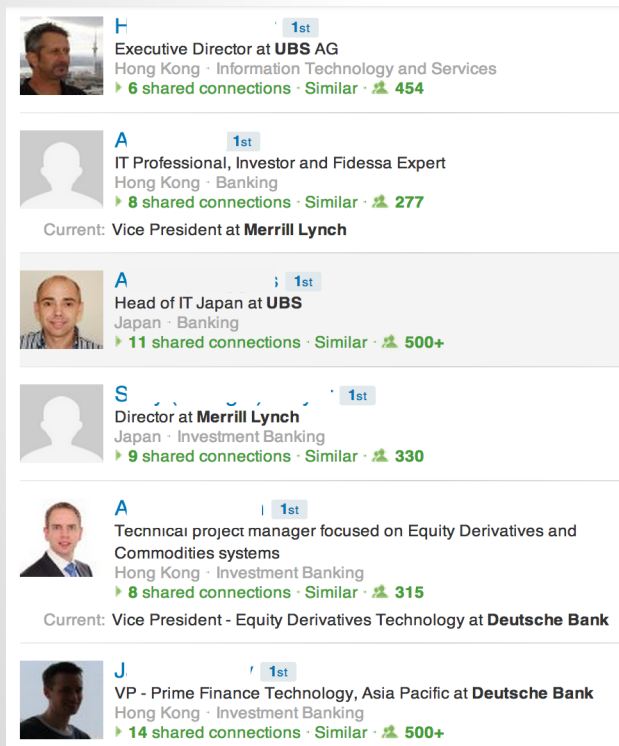
2nd Connections x Current Company: Merrill Lynch x Current Company: UBS x  
Current Company: Deutsche Bank x Reset

All same except the degree of connections

# Find Hidden LinkedIn Connections

Result: 57 vs 57 - 100% covered!

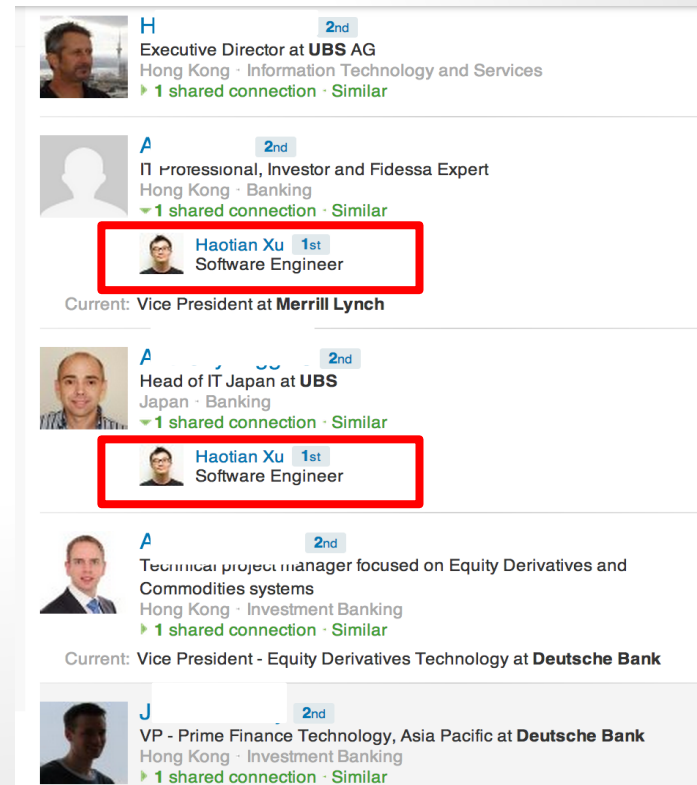
## Haotian (target)



Haotian Xu's LinkedIn profile is shown with five entries. Each entry includes a profile picture, name, title, location, and connection statistics. The entries are:

- H** Executive Director at **UBS AG**  
Hong Kong · Information Technology and Services  
▶ 6 shared connections · Similar · 👤 454
- A** IT Professional, Investor and Fidessa Expert  
Hong Kong · Banking  
▶ 8 shared connections · Similar · 👤 277  
Current: Vice President at **Merrill Lynch**
- A** Head of IT Japan at **UBS**  
Japan · Banking  
▶ 11 shared connections · Similar · 👤 500+
- S** Director at **Merrill Lynch**  
Japan · Investment Banking  
▶ 9 shared connections · Similar · 👤 330
- A** Technical project manager focused on Equity Derivatives and Commodities systems  
Hong Kong · Investment Banking  
▶ 8 shared connections · Similar · 👤 315  
Current: Vice President - Equity Derivatives Technology at **Deutsche Bank**
- J** VP - Prime Finance Technology, Asia Pacific at **Deutsche Bank**  
Hong Kong · Investment Banking  
▶ 14 shared connections · Similar · 👤 500+

## Ding Yi (attacker)



Ding Yi's LinkedIn profile is shown with five entries. Each entry includes a profile picture, name, title, location, and connection statistics. The entries are:

- H** Executive Director at **UBS AG**  
Hong Kong · Information Technology and Services  
▶ 1 shared connection · Similar
- A** IT Professional, Investor and Fidessa Expert  
Hong Kong · Banking  
▶ 1 shared connection · Similar  
**Haotian Xu** 1st Software Engineer (highlighted in a red box)  
Current: Vice President at **Merrill Lynch**
- A** Head of IT Japan at **UBS**  
Japan · Banking  
▶ 1 shared connection · Similar  
**Haotian Xu** 1st Software Engineer (highlighted in a red box)
- A** Technical project manager focused on Equity Derivatives and Commodities systems  
Hong Kong · Investment Banking  
▶ 1 shared connection · Similar  
Current: Vice President - Equity Derivatives Technology at **Deutsche Bank**
- J** VP - Prime Finance Technology, Asia Pacific at **Deutsche Bank**  
Hong Kong · Investment Banking  
▶ 1 shared connection · Similar

# Find Hidden LinkedIn Connections

- Tradeoff between the privacy protection and the utility
- Solution
  - Limit  $i^{\text{th}}$  degree search
  - Fuzzy name only available to trustworthy users
  - **Discussion - Other Ideas?**

# Linkedin iOS app case [5][6]

- Calendar viewing feature: opt-in
- Once enabled
  - Every calendar entry for the next 5 days
  - Meeting's title, organizer, attendees, time, notes
  - Both personal calendar entries and ones that may be for work
  - Information is being sent in **plain text**
- Dialog does NOT explicitly inform users about what information is being sent





# “Intro” iOS mail plugin case [7][8]

- An extension of iOS Mail Client - top bar
- CSS Tricks on *hover*: pseudo class
  - Cascading Style Sheets defines look and formatting for a document

David Buchanan  
To: rvohra@linkedin.com more...

Meeting?  
October 14, 2013 at 12:26 PM

Hi Mark! Tap to expand

Crosswise would love to work with you on an e-commerce project for one of our biggest clients.

I got a chance to look at some of the work you've done in the past. Great stuff!

Best,  
David

CEO at Crosswise  
Commona, Dynamics Inc.  
University of Oregon  
Austin, TX

LinkedIn  
David's Blog

How You're Connected  
Melissa Proctor, Kim Gupta, Dan Portello, S...

Summary  
Internet executive with over 10 years of experience. Particular expertise in user experience design, front-end engineering, and product management. Areas of interest include consumer internet and mobile, especially at massive scale.

<iframe src="https://...">

"Add connection" button

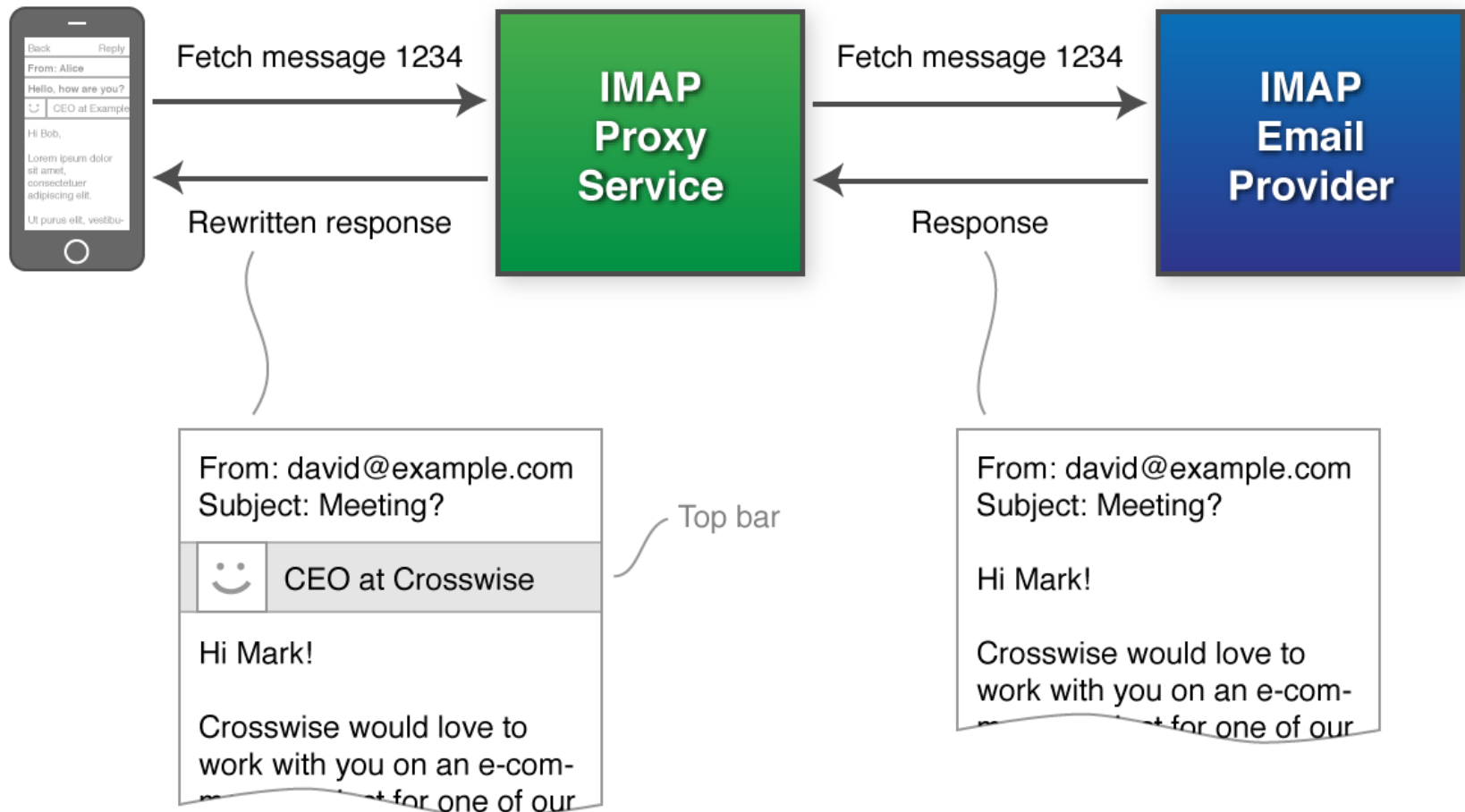
"Invited" status

"Connected" status

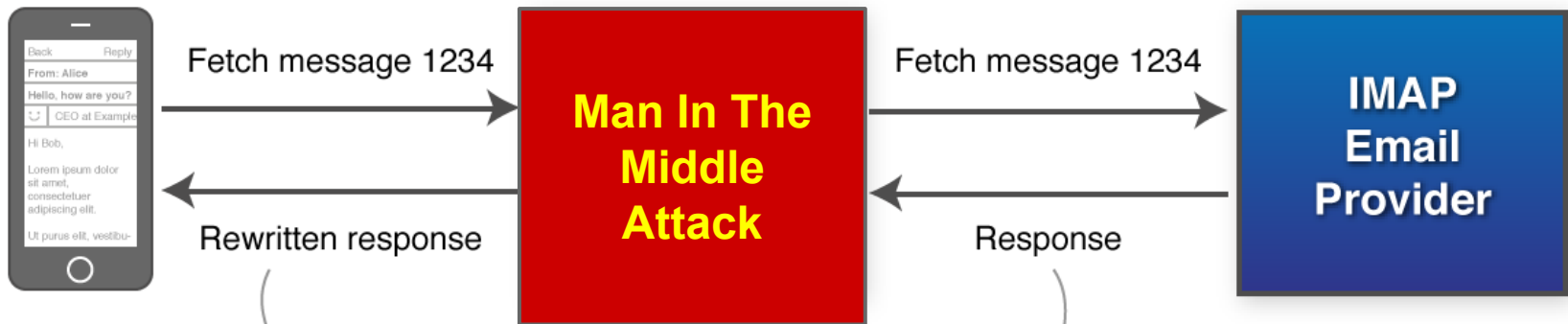
- **<iframe>** tricks

would love to work with  
e-commerce project for  
biggest clients.

# “Intro” iOS mail plugin case



# “Intro” iOS mail plugin case



From: david@example.com  
Subject: Meeting?

CEO at Crosswise

Hi Mark!

Crosswise would love to work with you on an e-commerce project for one of our

Top bar

From: david@example.com  
Subject: Meeting?

Hi Mark!

Crosswise would love to work with you on an e-commerce project for one of our

# “Intro” iOS mail plugin case <sup>[10]</sup>

- LinkedIn has access to plain text email
- Attorney-client privilege may be waived
- <iframe> is bad, possible clickjacking
- Changing device’s security profile
- Possible violation of company’s security policy
- LinkedIn’s Response - *Intro has 'the most secure implementation we believed possible'*

# Thoughts and Conclusion

- Not only LinkedIn - but also other SNS
- It's not what they say, but what they don't say on the privacy policy
- LinkedIn's social responsibility: should not facilitate identity theft, spam, stalking etc
- However...
  - Users voluntarily joined the network (or even pay for it) in exchange for the services
  - User accepted the agreement and terms
  - If it is truly sensitive, user should not put it onto SNS in the first place

# Reference

1. <http://press.linkedin.com/about>
2. <http://visual.ly/10-amazing-linkedin-statistics-2013>
3. [Link Prediction in Social Networks using Computationally Efficient Topological Features] by Michael Fire
4. Staddon, J. Finding "hidden" connections on LinkedIn, an argument for more pragmatic social network privacy. Proceedings of the 2nd ACM Workshop on Security and Artificial Intelligence (AI Sec 2009); 2009 November 9; Chicago, IL. NY: ACM; 2009; 11-14.
5. <http://thenextweb.com/insider/2012/06/06/linkedins-ios-app-collects-and-sends-names-emails-and-meeting-notes-from-your-calendar-back-in-plain-text>
6. <http://thenextweb.com/insider/2012/06/06/linkedin-updates-ios-app-following-data-transmission-concerns-but-insists-its-a-great-feature/>
7. <http://www.bishopfox.com/blog/2013/10/linkedin-intro/>
8. <http://www.theverge.com/2013/10/27/5034276/linkedin-says-intro-has-most-secure-implementation-possible>
9. <https://intro.linkedin.com/micro/privacy>