Social Snapshots

Digital Forensics for Online Social Networks

Markus Huber*†, Martin Mulazzani*, Manuel Leithner*, Sebastian Schrittwieser*, Gilbert Wondracek‡, Edgar Weippl*

*SBA Research
† Vienna PhD school of informatics
‡ Vienna University of Technology

ACSAC 2011, Orlando (FL)

1 Brief background
- Social Snapshots

2 Design
- Authentication
- Modules

3 Results and Evaluation
- Evaluation
- Results: information gathering

4 Concluding remarks

Brief Background

Online Social Networks (OSNs)

- Currently largest and fasted growing web services
- Personal data of hundreds of million people
- Facebook, LinkedIn, XING, etc.

1. **Replace traditional means of digital storage, sharing, and communication**
   - Need for novel digital forensics data collection methods
2. **Research on OSNs security issues**
   - Empirical studies [1, 2, 4, 3, 5] depend on initial data gathering

State-of-the-art information gathering methods

- Extraction of sensitive information poses non-trivial challenge
- Simple web crawlers (libwww etc.)
- Number of shortcomings
  1. High network traffic
  2. Additional or hidden data
  3. Maintainability

Main contributions

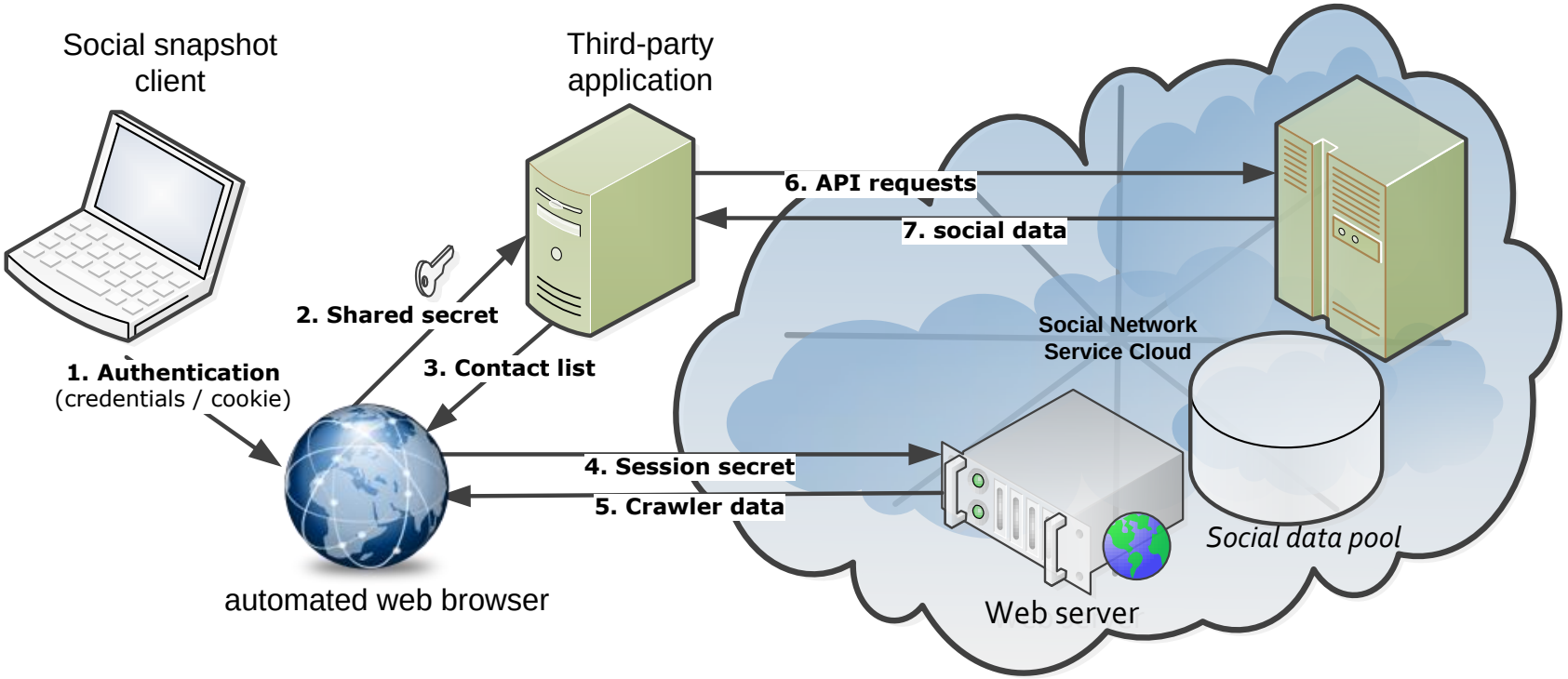
- Novel techniques to gather OSNs data
  - Hybrid approach: Third-party application + crawling
- Social Snapshot prototype for Facebook
  - Core framework released as open-source software
- First experimental evaluation
  - Based on Facebook

Design

Design goals

- Simulate average user
  - Limit network-traffic, gathering-duration
  - Standard Web Browser
- Collection of meta-data
  - Rely on rich data available through APIs
- Open-source software (OSS)

Social Snapshot Framework



How-to gather the initial authentication token

- Consent
  - Easiest case, preferred method for research
- Hijack social networking sessions
  - WiFi, LAN
- Extraction from forensic image
  - Stored authentication cookies from seized hard-drives

Social Snapshot Framework modules

- Social snapshot client
- Automated web browser
- Third-party social snapshot application
- Hijack
- Digital image forensics
- Analysis

Results and Evaluation

Evaluation based on Facebook

- Evalation based on Facebook
  - At the time of writing: largest online social network
  - Support for third-party applications
  - Graph API enables access great majority of account content

Element	Download	social snapshot
Contact details	–	✓ Crawler
News feed	–	✓ Graph API
Checkins	–	✓ Graph API
Photo Tags	–	✓ Graph API
Video Tags	–	✓ Graph API
Friends	name only[a]	✓ Graph API
Likes	name only[a]	✓ Graph API
Movies	name only[a]	✓ Graph API
Music	name only[a]	✓ Graph API
Books	name only[a]	✓ Graph API
Groups	name only[a]	✓ Graph API
Profile feed (Wall)	limited[b]	✓ Graph API
Photo Albums	limited[b]	✓ Graph API
Video Uploads	limited[b]	✓ Graph API
Messages	limited[b]	✓ Graph API

Graph API Result example


```
{"id":"12345678","name":"John Doe",
"first_name":"John","last_name":"Doe",
"link":"http:\/\/www.facebook.com\/johndoe",
"username":"johndoe","birthday":"04\/01\/1975",
"hometown":{"id":"","name":null},
"quotes":"social snapshot your account!.\n",
"gender":"male","email":"johndoe@example.com",
"timezone":2,"locale":"en_US","verified":true,
"updated_time":"2011-05-15T13:05:19+0000"}
```

Challenges

- Hijack module
  - No direct support to set cookies in Selenium *(patch for server)*
- Graph API SDK
  - Too slow *(modified original lib)*
- Crawling of contact details
  - List of friends *(via API)*
  - Plaintext emails replaced with images *(image generation script)*
  - Emails get removed completely *(address.yahoo.com)*
- Recruiting test subjects
  - Challenging *(25 test subjects)*

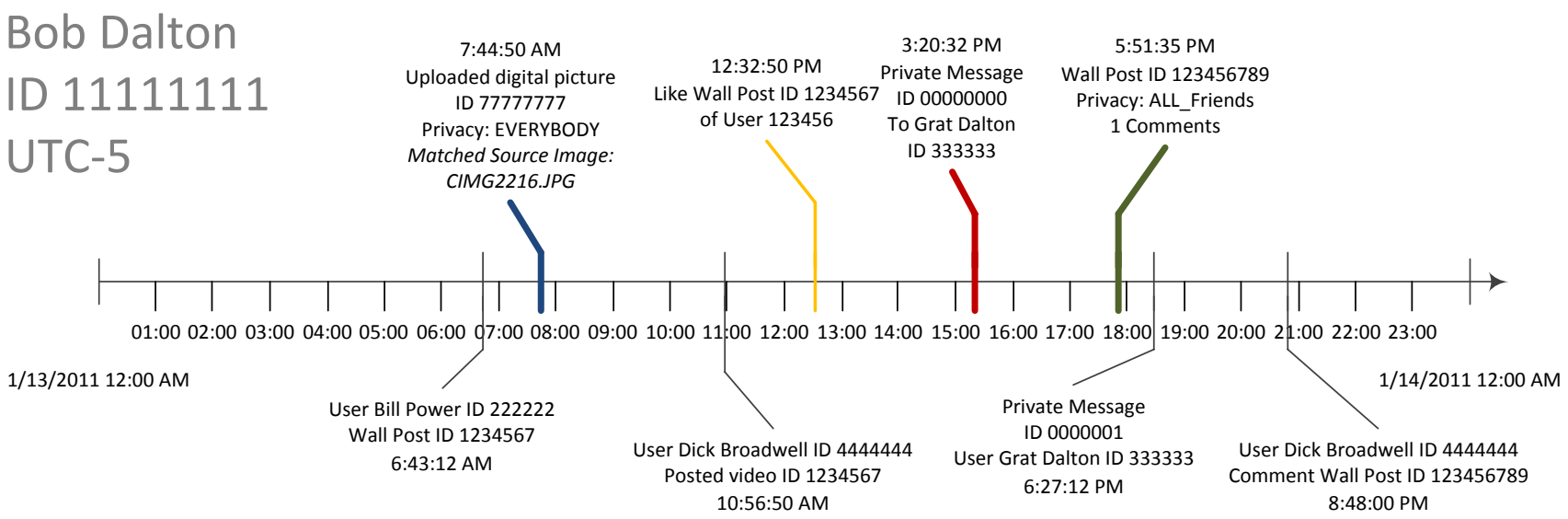
Results: information gathering

- Run-Time
  - Third-party application 12.79min on average
  - Crawler 14min on average
- Fetched elements
  - Third-party required **9802** API requests on average
  - Crawler processed **238** friend profiles on average
  - After **162** plaintext email addresses we had to solve **85** email addresses with OCR

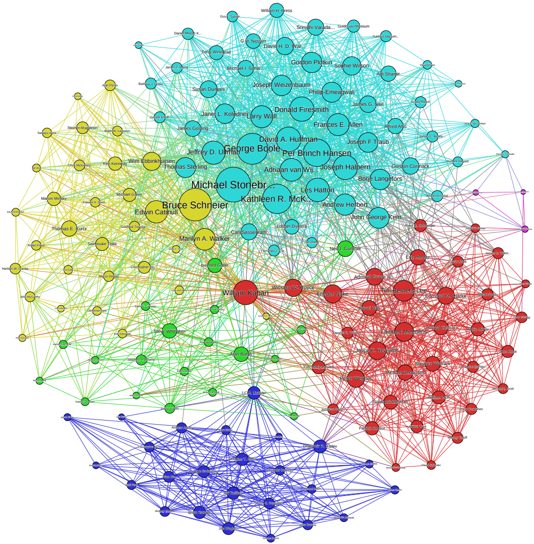
Results: information gathering

Video/demo.avi

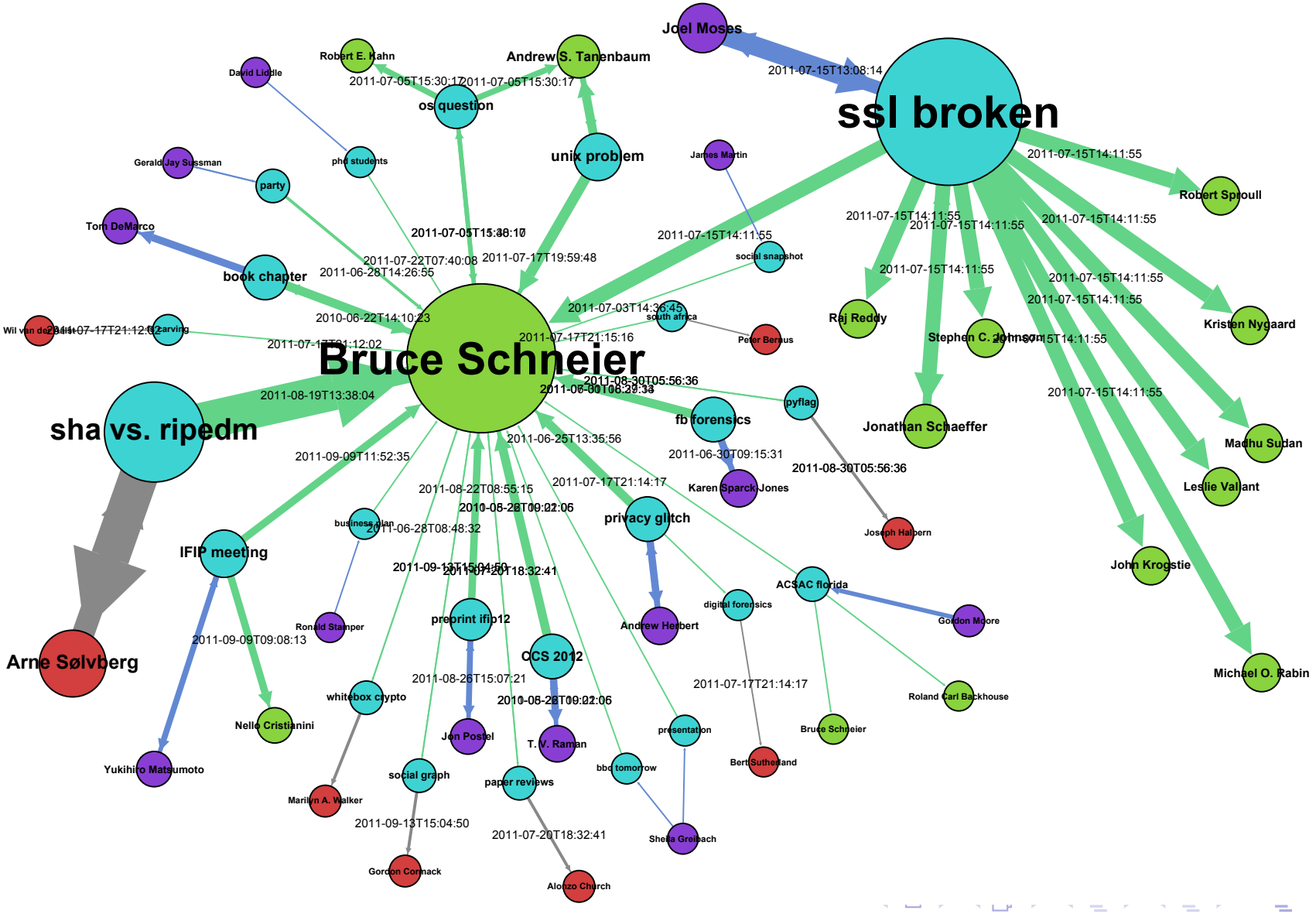
Results: information gathering



Results: information gathering



Results: information gathering



Concluding remarks

- Social Snapshot tool extracts Facebook data in less than 15 minutes
- Malicious social snapshots
  - FiTM attacks[3]
- Privacy surveys
- Data liberation

Thank you for your time!

Questions?

mhuber@sba-research.org

http://socialsnapshot.nysos.net

Participate in our survey and get your social snapshot:

http://is.gd/snapshotsurvey

L. Bilge, T. Strufe, D. Balzarotti, and E. Kirda.
All your contacts are belong to us: automated identity theft attacks on social networks.
In *Proceedings of the 18th international conference on World wide web*, pages 551–560. ACM, 2009.

H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Zhao.
Detecting and characterizing social spam campaigns.
In *Proceedings of the 10th annual conference on Internet measurement*, pages 35–47. ACM, 2010.

M. Huber, M. Mulazzani, E. Weippl, G. Kitzler, and S. Goluch.
Friend-in-the-middle attacks: Exploiting social networking sites for spam.
Internet Computing, 2011.

T. Jagatic, N. Johnson, M. Jakobsson, and F. Menczer.
Social phishing.
Communications of the ACM, 50(10):94–100, 2007.

G. Wondracek, T. Holz, E. Kirda, and C. Kruegel.
A Practical Attack to De-Anonymize Social Network Users.
In *Proceedings of the IEEE Symposium on Security and Privacy*, 2010.