



a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA

Laboratory for Education and Research in Secure Systems Engineering (LERSSE)

Department of Electrical & Computer Engineering

outline

- why OSNs?
- rewards and challenges of research in OSN
- current research directions
 - de-anonymization
 - privacy (game)
 - Sybil & compromised account detection/resistance

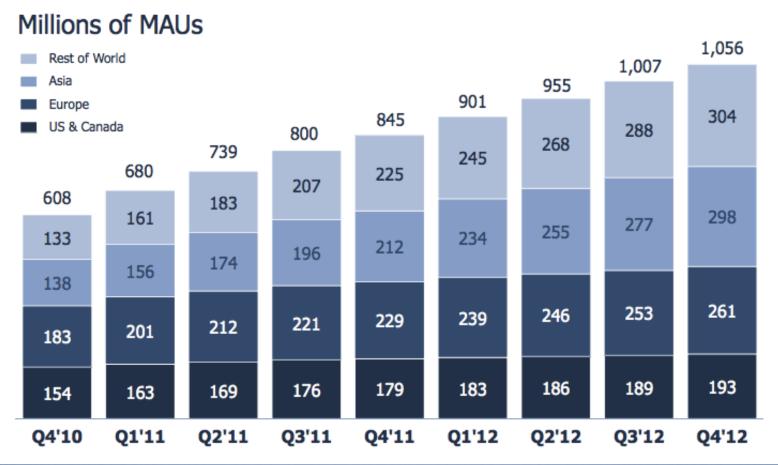


why OSNs?: multitude

Site	Traffic Rank	Users (M)	Country
Windows Live	4	120	USA
Facebook	4	175	USA
MySpace	7	250	USA
Hi5	17	60	USA
SkyRock	43	13	France
Friendster	45	95	USA
NetLog	71	35	Belguim
Tagged	75	70	USA
Orkut	83	67	USA
LiveJournal	85	18	Russia
Bebo	119	40	USA
PerfSpot	124	20	USA
meinVZ	156	12	Germany
Multiply	161	12	USA
Badoo	168	19	UK
Sonico	183	33	Argentina
Ning	187	1	USA
CyWorld	315	20	South Korea
Xanga	346	40	USA
MyYearbook	406	15	USA

why OSNs?: sheer scale

Monthly Active Users (MAUs)



Please see Facebook's Form 10-K for the year ended December 31, 2012 for definitions of user activity used to determine the number of our MAUs, DAUs and mobile MAUs. The number of MAUs, DAUs, and mobile MAUs do not include Instagram users unless such users would otherwise qualify as MAUs, DAUs, and mobile MAUs based on activity that is shared back to Facebook.

In June 2012, we discovered an error in the algorithm we used to estimate the geographic location of our users that affected our attribution of certain user locations for the first quarter of 2012. The first quarter of 2012 user metrics reflect a reclassification to more correctly attribute users by geographic region.

facebook.

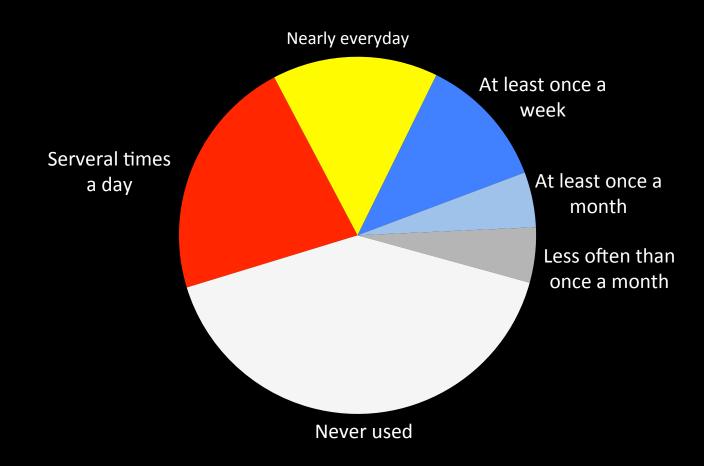
why OSNs?: this is where users are!

- 20% of US page-views are on Facebook [1]
- Each Facebook user spends on average 15 hours and 33 minutes a month on the site [2]
- Twitter is handling 1.6B queries per day [2]

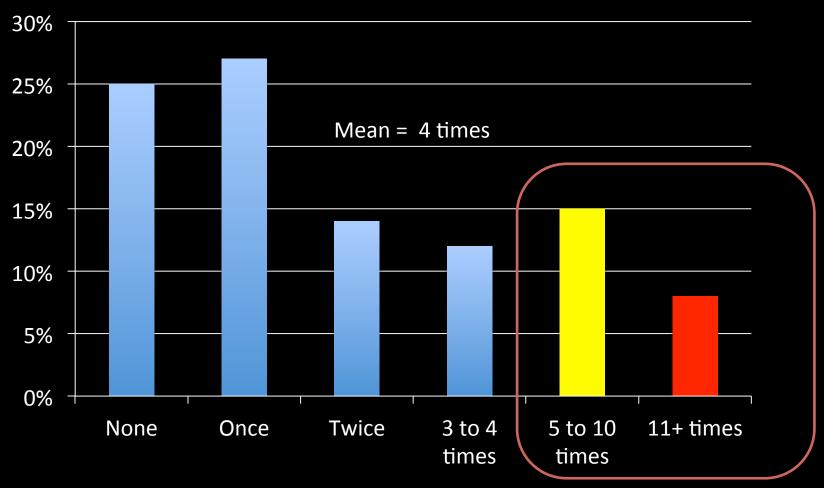
^{1.} Jennifer Beese "Facebook Accounts for 20 Percent of U.S. Pageviews" http://sproutsocial.com/insights/2012/02/facebook-stats

^{2.} Jeff Bullas "20 Stunning Social Media Statistics Plus Infographic," jeffbullas.com, 2011-Sep-02

nearly 2/3 of OSN users use them daily



In the last 25 hours, approximately how many times did you check your Facebook account?



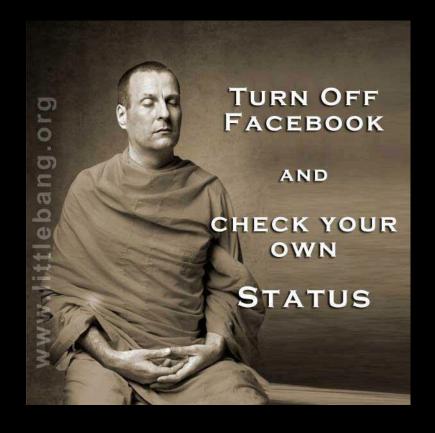
"The Social Habit" Edison Research and Arbitron, 2012-June

what do users do there?

 social connection shared identities Photographs OUI content COL sat social invest COL COL social netwo What's your Smi status updat con Status is c Update? boti con

.Joinson, A. "Looking at", "looking up" or "keeping up with" people?: Motives and use of Facebook. In CHI 2008, ACM (2008), pp. 1027-1036.





why OSNs?: reach out real world



Obama raised \$690m online in 2012.

50m 'likes'

why OSNs?: mobilize real world

Arab Spring in 2011

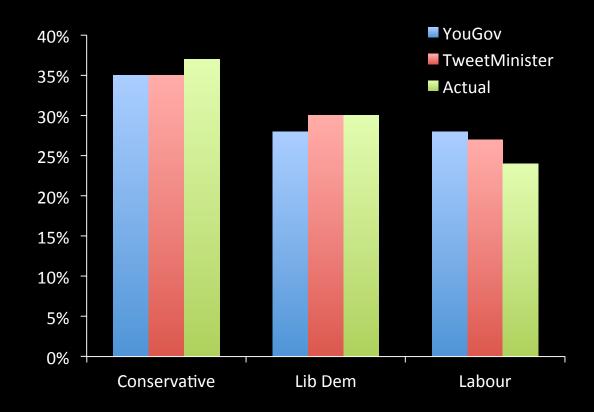


Photo credit: Peter Macdiarmid, Getty Images



Photo credit: Steve Crisp, Reuters

why OSNs?: reflect real world

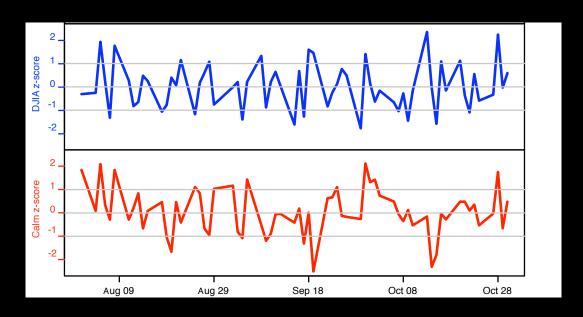


predicting the future: UK General Election 2010

Anthony Wells, GovMonitor, 6 May 2010, http://www.thegovmonitor.com/world-news/britain/uk-election-2010-final-polls-30081.html Jemima Kiss, The Guardian, 13 May 2010, http://www.guardian.co.uk/media/pda/2010/may/13/twitter-tweetminster-election

why OSNs?: predict real world

Twitter mood (Calm) predicts Dow Jones Industrial Average (DJIA)



Bollrn et al. "Twitter mood predicts the stock market" J. Comp. Sc., March, 2011.

why OSN?

- multitude
- sheer scale
- this is where users are!
- reach out real world
- mobilize real world
- reflect real world
- predict real world

WHAT MAKES OSNs ATTRACTIVE FOR RESEARCH?

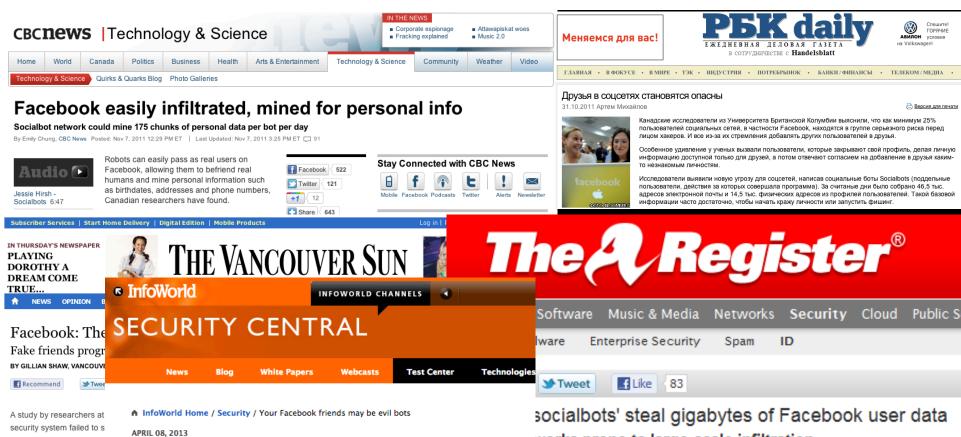
what's attractive?: diverse stakeholders and actors

- users
- OSN operator
- advertisers
- OSN application developers
- user's employers, ensures, etc.
- law enforcement, intelligence agencies, and other government organizations
- stalkers, investigators, users' nosy colleagues and neighbors

what's attractive?

also ...

- new phenomena
- volatile
- socio-technical systems
- new threats, vulnerabilities, defenses
- most of us are OSN users
 - easier to recruit study participants
 - relevance



generated fake Facebook



2 November 2011 Last update

Socialbots u Facebook da

Researchers have demons technique capable of stea information from Faceboo

Using 'socialbots', computer mimic real Facebook profiles were able to harvest vast qu data

Socialbots are increasingly t internet criminals and are be on the internet for as little as \$29 (£18).

Your Facebook friends may be evil bots

Computer scientists have unleashed hordes of humanlike social bots to infiltrate Facebook -- and they're awfully effective

By Eagle Gamma | InfoWorld

Print | 50 6 Comments

















How safe is your online social network? Not very, as it turns out. Your friends may not even be human, but rather bots siphoning off your data and influencing your decisions with convincing yet programmed points of view.

A team of computer researchers at the Department of Electrical and Computer Engineering at the University of British Columbia has found that hordes of social bots could not only spell disaster for large online

то ассерт тие таке тпепо



Credit: Palto/iStockphoto



works prone to large-scale infiltration



的位置: 比特网>安全>正文

黑客软件成功窃取Facebook海量用户信息

2011-11-08 21:38 卡饭资讯 lu

字号: A+ A-

加拿大温哥华不列颠哥伦比亚大学四名研究员开发出黑客软件"社交机器人(socialbots)"。成 功窃取了Facebook海量用户数据。在传统僵尸网络中, 黑客用病毒感染电脑后进行远程控制, 窃取受害者电脑数据,或者使用被感染电脑发送垃圾信息实施更多攻击。"社交机器人"则完全 模拟Facebook真实用户操作、自动设定用户名称和头像、随机发送好友申请。

研究员将102个社交机器人用于实验,由"主机器人"向其他机器人发送命令。机器人每天发 送25个好友申请,实验为期8周多,机器人总共向8570个Facebook账号发送了好友申请,3055位用 户接受申请。研究员发现好友人数多的Facebook用户更易接受假好友申请。

what's attractive? summary

- diverse stakeholders and actors
- new phenomena
- volatile
- socio-technical systems
- new threats, vulnerabilities, defenses
- easy to recruit study participants
- relevant

WHAT MAKES OSN RESEARCH CHALLENGING?

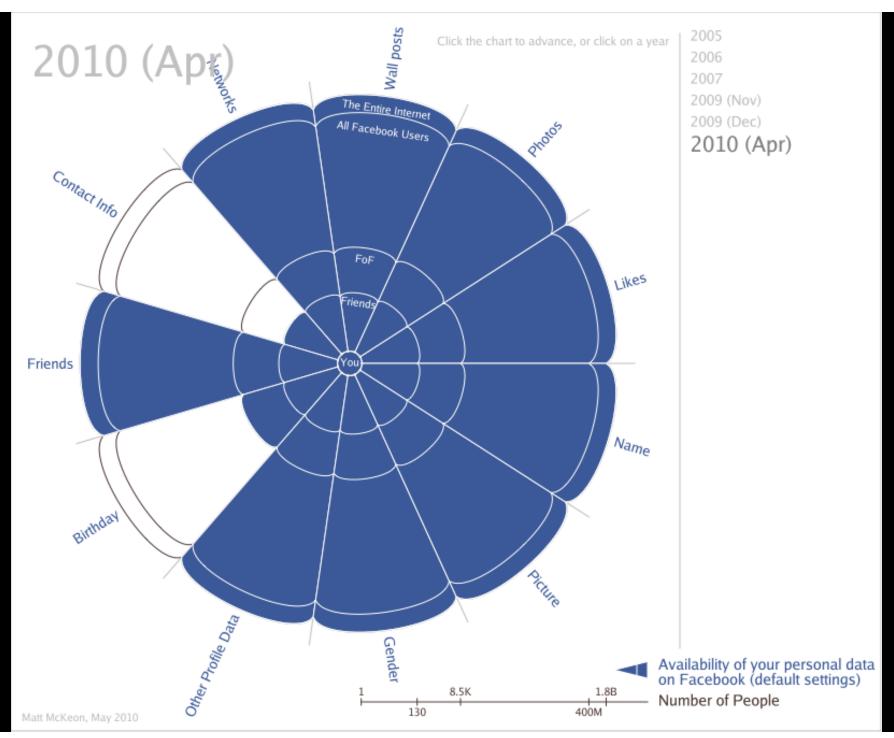
what's challenging?

- overcrowded by researchers
- access to data becoming difficult
- hard to evaluate vulnerabilities/defenses

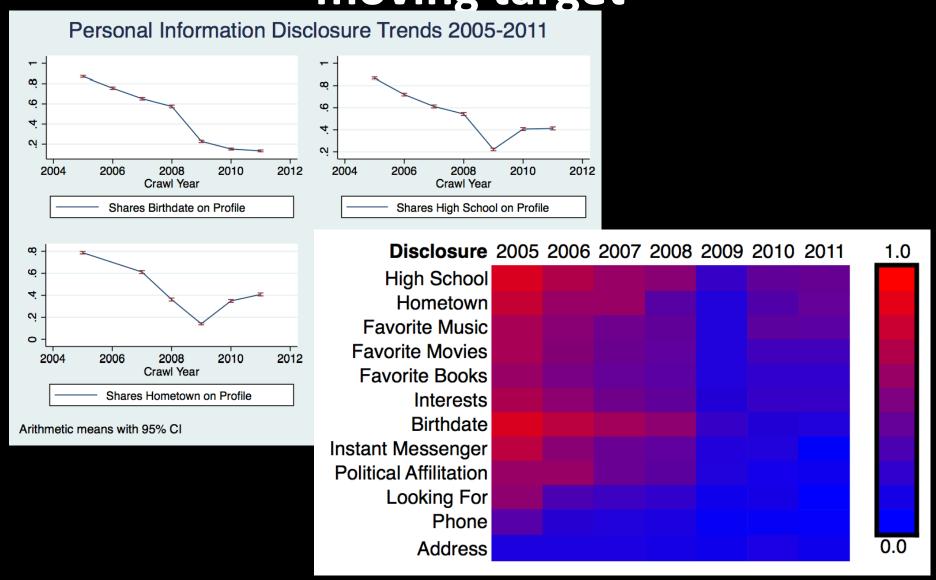




just in January-May 2013



what's challenging?: moving target



Fred Stutzman, Ralph Gross, Alessandro Acquisti, "Silent Listeners: The Evolution of Privacy and Disclosure on Facebook," Journal of Privacy and Confidentiality (2012) 4, Number 2, pp. 7–41.

what's challenging?: OSN operators protective

facebook

We are reluctant to take legal action against UBC, its researchers or students engaged in legitimate academic projects. However, this is not the first time UBC has ignored Facebook's

Facebook must insist that UBC and its researchers abide by Facebook's terms and the law. Additionally, given the apparent ongoing and knowing disregard of Facebook's terms, the law and UBC ethical obligations, we request that your offices: (1) ensure that UBC researchers cease and desist any and all unauthorized access to Facebook's site and systems; (2) return to Facebook all illegally harvested user data obtained by UBC researchers and certify destruction of all copies that remain in UBC's possession; (3) provide an accounting of all research activities involving Facebook and its users; (4) suspend any ongoing Facebook-related research unless and until Facebook provides consent; (5) explain the process by which UBC approved this particular study; and (6) preserve all materials that refer or relate to the UBC's approval, or lack thereof, for studies involving Facebook and/or its users.

summary of challenges

- overcrowded by researchers
- access to data becoming difficult
- hard to evaluate vulnerabilities/defenses
- moving target
- OSN operators protective

RESEARCH IN OSN SECURITY & PRIVACY

research directions

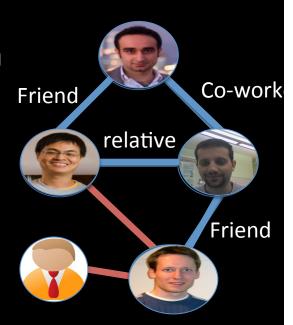
- de-anonymization
- privacy (game)
- sybil detection/resistance

DE-ANONYMIZATION OF/WITH OSNs

social network data anonymization?

why?

- academic and government research
- advertising
- third-party applications
- aggregation
- how?
 - remove node or edge attributes
 - inject random noise

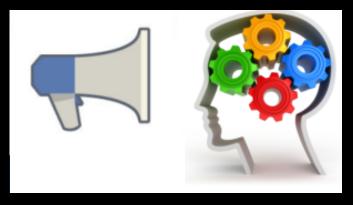


F. Beato, M. Conti, and B. Preneel, "Friend in the Middle (FiM): Tackling De-Anonymization in Social Networks," IEEE International Workshop on SEcurity and SOCial Networking, 6 p., 2013.

threat agents in de-anonymization attacks



large-scale collection of detailed information on individuals



abusive marketing aimed at specific individuals



craft a highly individualized, believable message

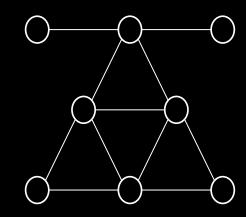


stalkers
investigators
nosy
colleagues
employers
neighbors

recognize the victim's node in the anonymized network and to learn sensitive information

de-anonymization

- active [1]
 - "mark" regions of the graph with injected nodes (Sybils) and/or edges



- costly on large scale
- passive [2]
 - use "auxiliary" network to re-identify nodes
 - self-reinforcing: seed population increases

^[1] Lars Backstrom, Cynthia Dwork, and Jon Kleinberg, "Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography" In international conference on World Wide Web (WWW), pp. 181-190, 2007.

^[2] Narayanan, Arvind, and Vitaly Shmatikov. "De-anonymizing social networks," In IEEE Symp. on Sec. & Privacy, pp. 173-187. IEEE, 2009.

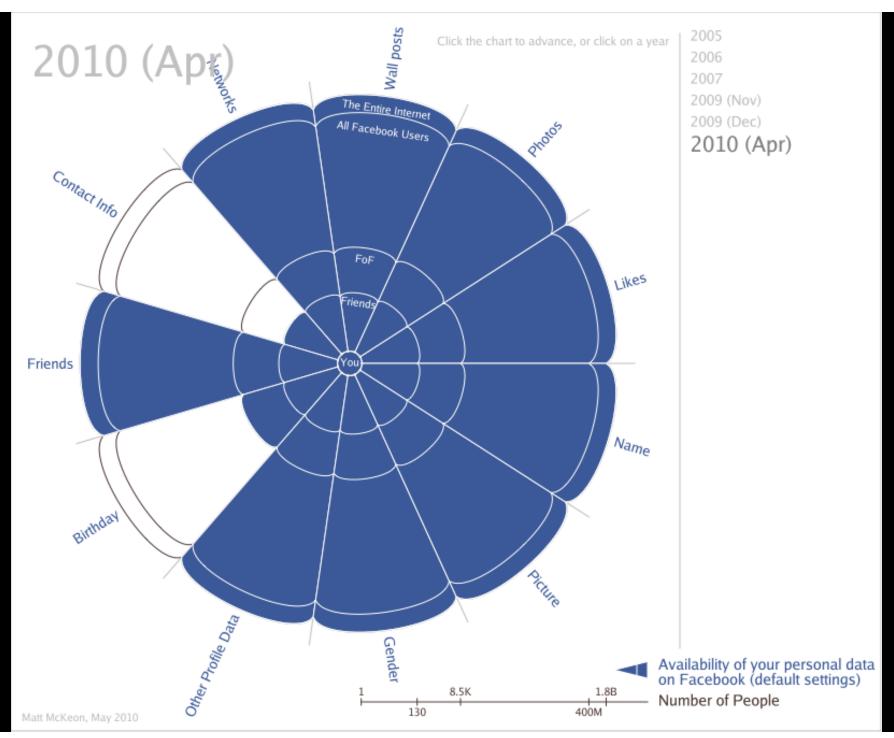


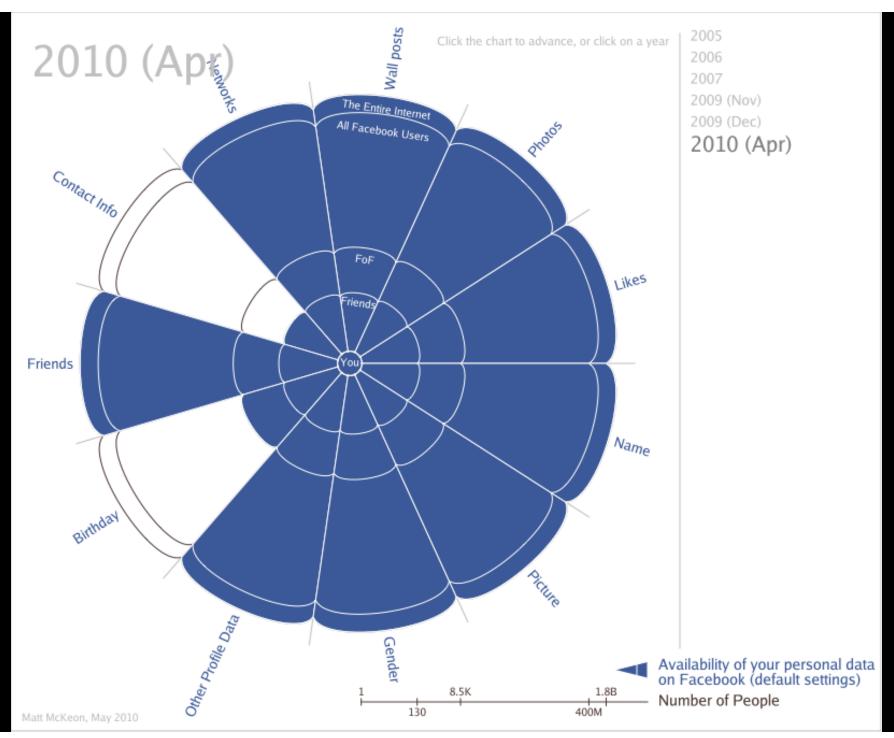
operators collect lots of personal data ...



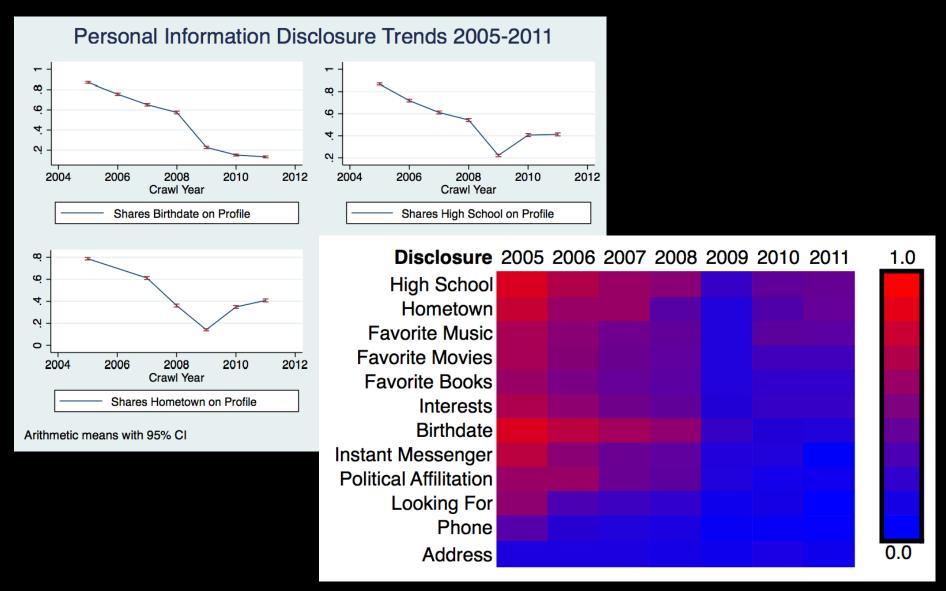
... and then make it widely visible

Visibility Level	Default	Optional	Unavailable
Public Internet	41%	-	59%
All site users	48%	28%	24%
Sub-networks only	7%	17%	76%
Friends of friends	-	24%	76%
Friends only	3%	79%	17%





how do users (re)act?

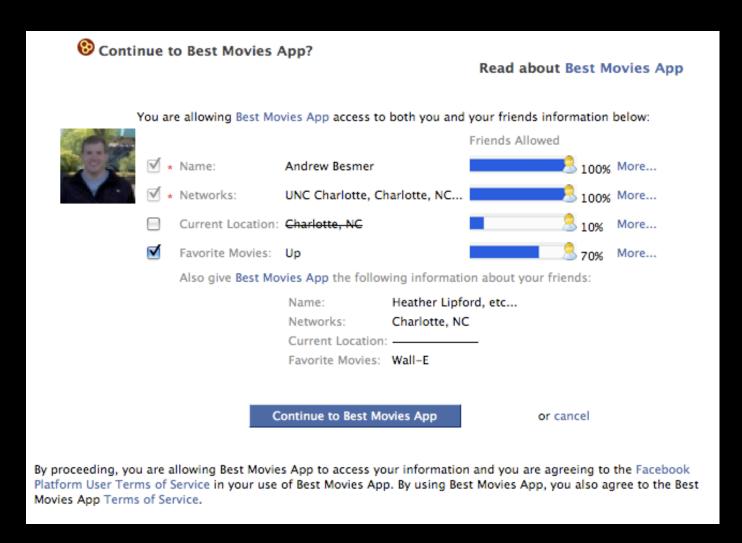


Fred Stutzman, Ralph Gross, Alessandro Acquisti, "Silent Listeners: The Evolution of Privacy and Disclosure on Facebook," Journal of Privacy and Confidentiality (2012) 4, Number 2, pp. 7–41.

what makes young users to "lock" data?

- 1. no marginal gain for maintenance of an open profile, when user's network becomes "large"
- 2. <u>expectancy violations by weak-ties</u> generate privacy concerns among females
- 3. <u>engaging in conversational management of privacy</u> leads to customizing controls

improving models and UIs



Andrew Besmer, Heather Richter Lipford, Mohamed Shehab, and Gorrell Cheek, "Social applications: exploring a more secure framework," In Proceedings of the 5th Symposium on Usable Privacy and Security (SOUPS '09). Article 2, 10 p.

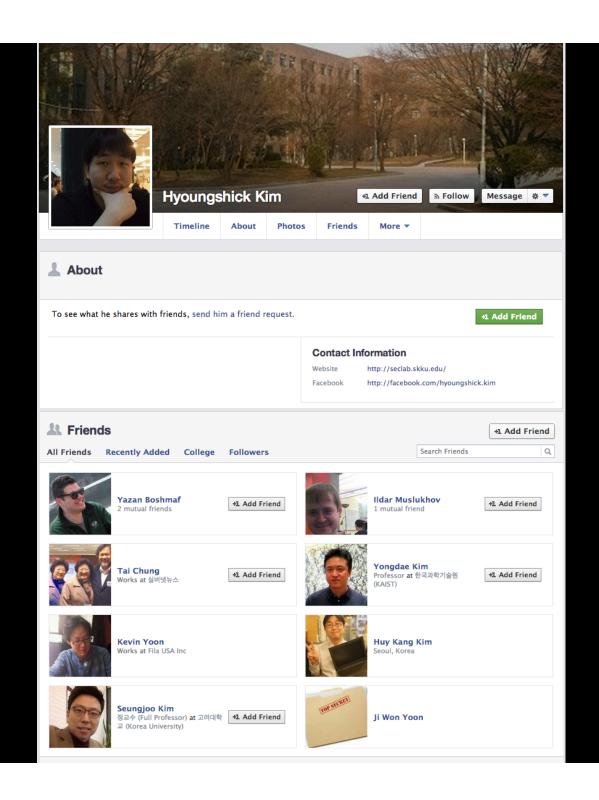
privacy communication game

- optimizes OSN interaction with each user group
 - pragmatic majority
 - claims to be interested in privacy
 - forgets about privacy when given an attractive service or monetary rewards
 - more assurance of privacy can make them less comfortable than simply ignoring privacy
 - privacy fundamentalists
 - care deeply about privacy, and
 - may actively investigate a site and complain to non-fundamentalists
- minimize the concerns of the fundamentalists while simultaneously minimizing the awareness of privacy for the pragmatic majority
- poor privacy may be a rational choice for operators

research directions

- de-anonymization
- privacy (game)
- sybil detection/resistance
- detection of compromised accounts

SYBIL DETECTION/RESISTANCE IN OSNs





Zhi Yang, Christo Wilson, Xiao Wang, Tingting Gao, Ben Y. Zhao, and Yafei Dai, "Uncovering social network sybils in the wild," In Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference (IMC '11).

Sybils can be helpful ...



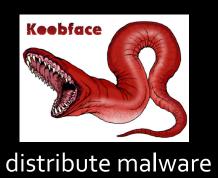
ECE, Olin College [1]



Web Ecology Project [2]

^[2] Socialbots competition: http://www.webecologyproject.org/category/competition/

... or dangerous



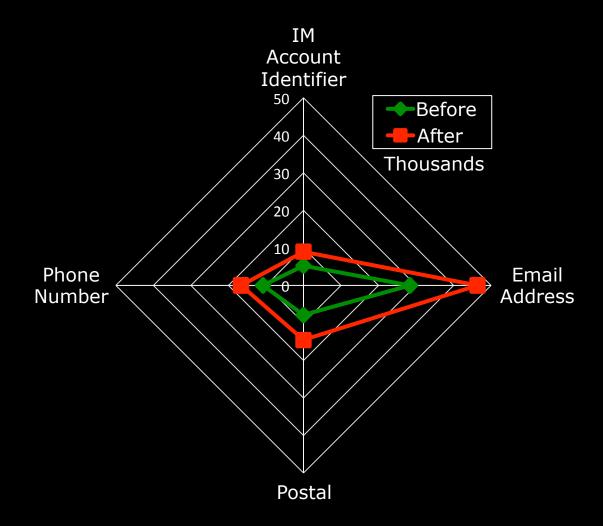


spread misinformation



collect data

can collect personal data



Y. Boshmaf, I. Muslukhov, K. Beznosov, M. Ripeanu, "Design and analysis of a social botnet," Elsevier Computer Networks, Special Issue of Botnet Design and Takedown, February 2013, pp. 556-578.

most importantly: can erode trust in ecosystem



Facebook Applications



Facebook Connect

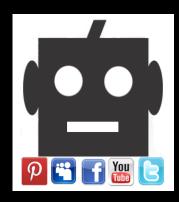
socialbots



Software



Social media account





ECE, Olin College



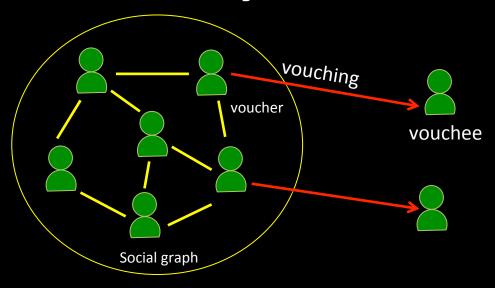
The Web Ecology Project

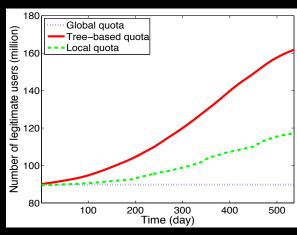
- [1] Dan Misener. Rise of the socialbots: They could be influencing you online. CBC News, March 2011.
- [2] Hwang et al. Socialbots: voices from the fronts. ACM Interactions 19, 2 (March 2012), 38-45.

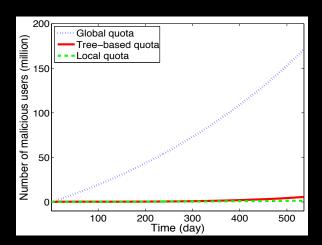
approaches to reducing sybils in OSNs

- admit into OSN carefully
 - increase trust slowly, by observing actions
 - hurts growth of OSNs, turns users away
- detect (and disable) sybils
 - give full trust right away
 - analyze graph or individual accounts
 - graph-based detection
 - classification based on account "behavior"
 - challenge suspects
- make it hard for sybils to infiltrate the OSN
 - do users care?
 - how can they make better decisions?

innocent by association







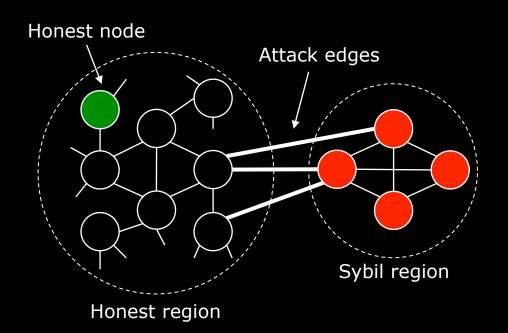
legitimate user growth

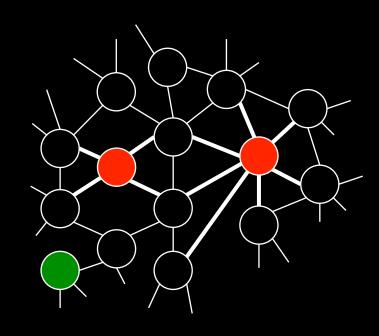
malicious user growth

approaches to reducing sybils in OSNs

- admit into OSN carefully
 - increase trust slowly, by observing actions
 - hurts growth of OSNs, turns users away
- detect (and disable) sybils
 - give full trust right away
 - analyze graph or individual accounts
 - graph-based detection
 - classification based on account "behavior"
 - challenge suspects
- make it hard for sybils to infiltrate the OSN
 - do users care?
 - how can they make better decisions?

Graph-theoretic Defense Techniques

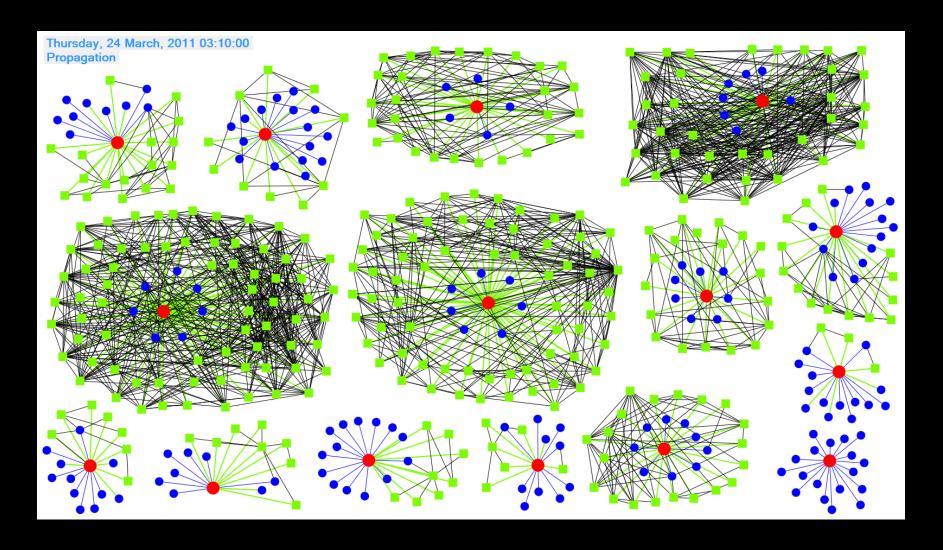




Sybil detection via social networks [1]: SybilRank, SybilLimit, SybilGuard, SybilInfer, GateKeeper in reality it could be like this [2]

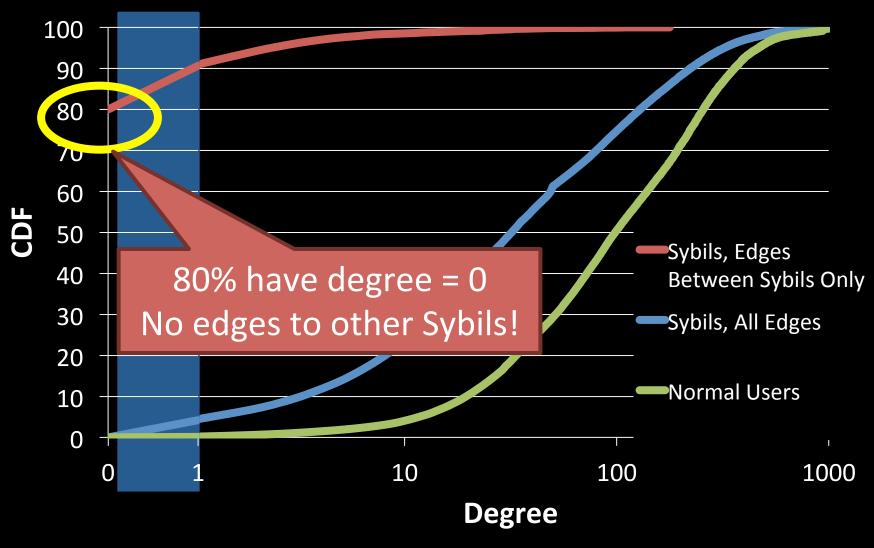
[1] Haifeng Yu. 2011. Sybil defenses via social networks: a tutorial and survey. SIGACT News 42, 3 (October 2011), pp. 80-101. [2] Boshmaf et al. Graph-based Sybil detection in social and information systems. To appear in the Proceedings of IEEE/ACM ASONAM, Niagara Falls, ON, Canada (August 2013).

a counter-example



Y. Boshmaf, I. Muslukhov, K. Beznosov, M. Ripeanu, Design and analysis of a social botnet. Elsevier Computer Networks – Special Issue of Botnet Design and Takedown, February 2013, pp. 556-578.

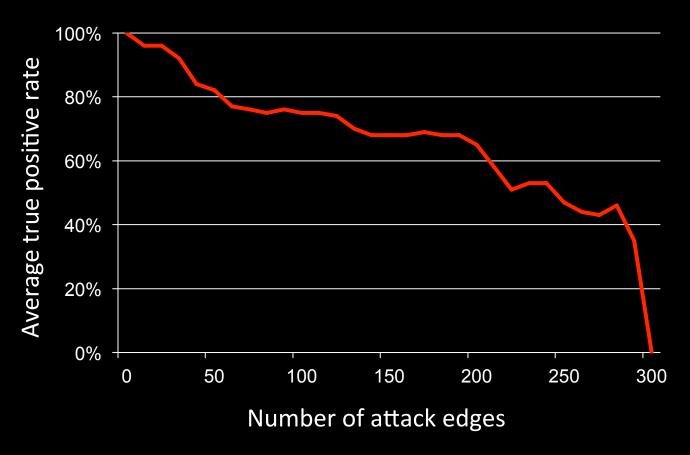
do Sybils form connected components?



Zhi Yang, Christo Wilson, Xiao Wang, Tingting Gao, Ben Y. Zhao, and Yafei Dai, "Uncovering social network sybils in the wild," In Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference (IMC '11).

from 100% TPR to 0% in 2 weeks

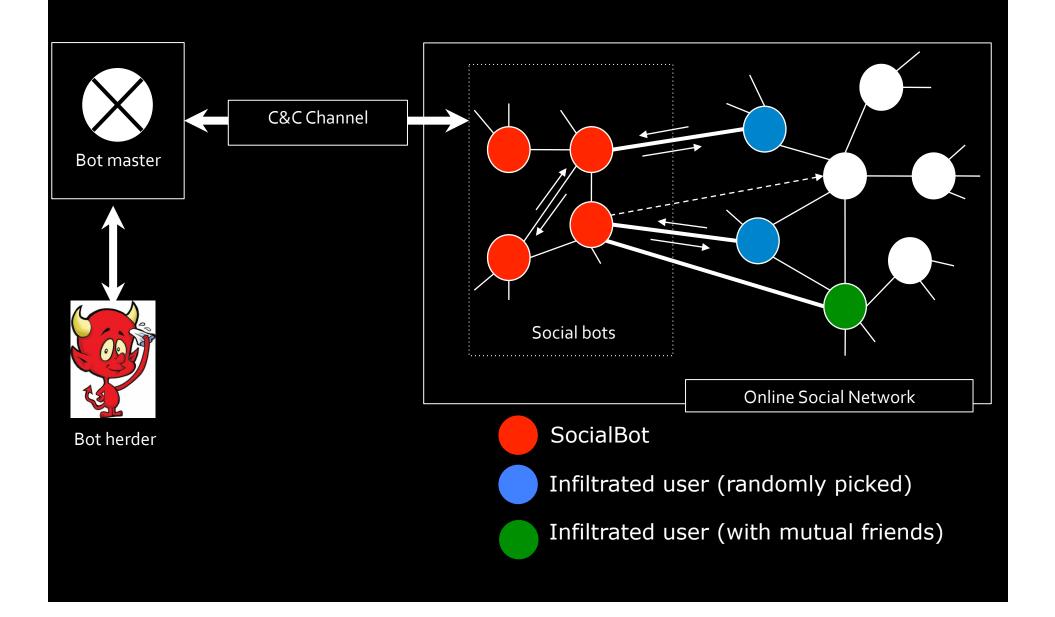
Using state-of-the-art local community detection algorithm to detect Sybils during the first two weeks



Yazan Boshmaf, Konstantin Beznosov, Matei Ripeanu, "Graph-based Sybil detection in social and information systems," in the Proceedings of IEEE/ACM ASONAM, Niagara Falls, ON, Canada (August 2013).

HOW FEASIBLE IS THE RISK OF SYBILS?

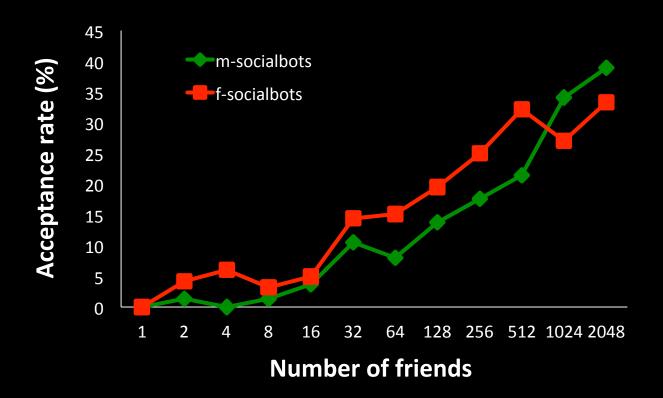
a more real example: a social botnet





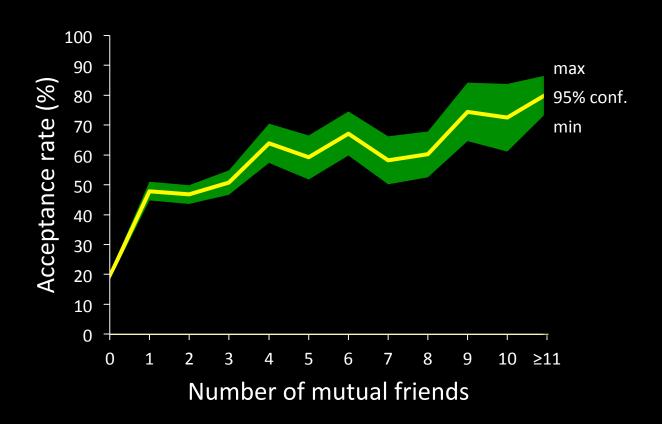


more friends, more Sybils



Y. Boshmaf, I. Muslukhov, K. Beznosov, M. Ripeanu, "Design and analysis of a social botnet," Elsevier Computer Networks – Special Issue of Botnet Design and Takedown, February 2013, pp. 556-578.

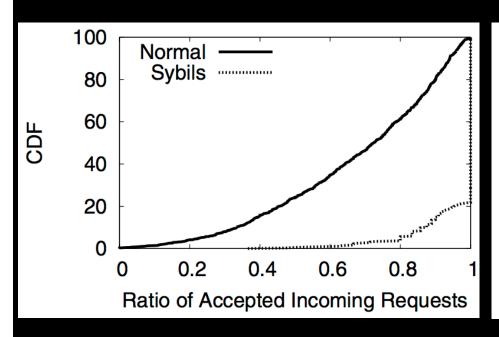
mutual friends matter

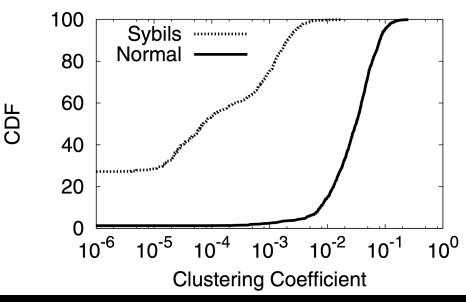


Y. Boshmaf, I. Muslukhov, K. Beznosov, M. Ripeanu, Design and analysis of a social botnet. Elsevier Computer Networks – Special Issue of Botnet Design and Takedown, February 2013, pp. 556-578.

possible Sybil indicators

- friend request frequency
- outgoing friend requests accepted





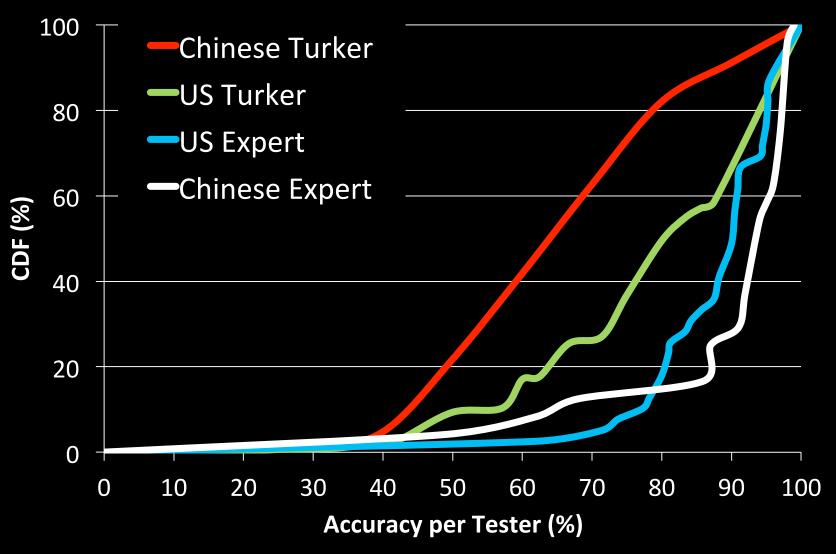
Zhi Yang, Christo Wilson, Xiao Wang, Tingting Gao, Ben Y. Zhao, and Yafei Dai, "Uncovering social network sybils in the wild," In Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference (IMC '11).

Gang Wang, Manish Mohanlal, Christo Wilson, Xiao Wang, Miriam Metzger, Haitao Zheng and Ben Y. Zhao, "Social Turing Tests: Crowdsourcing Sybil Detection," NDSS '13.

Sybil.Detector

- Previous <u>1</u> 2 <u>38</u>	39 40 41 42 43 44 45 46 49 50 Next →
The below profile is	Save changes If fake, mark suspicious content (multiple choice
Real Account	Profile Info
Fake Account	■ Wall
	Photos
nfo Wall Photos)
	Rachel Thompson Morked at Victoria Secret Studied at Harvard University Lives in New York, New York From Paris, France Work and Education
Wall Info	
Wall Info Photos Friends Subscriptions (32) Subscribers (117)	Morked at Victoria Secret

experts detect Sybils much better



Gang Wang, Manish Mohanlal, Christo Wilson, Xiao Wang, Miriam Metzger, Haitao Zheng and Ben Y. Zhao, "Social Turing Tests: Crowdsourcing Sybil Detection," NDSS '13.

approaches to reducing sybils in OSNs

- admit into OSN carefully
 - increase trust slowly, by observing actions
 - hurts growth of OSNs, turns users away
- detect (and disable) sybils
 - give full trust right away
 - analyze graph or individual accounts
 - graph-based detection
 - classification based on account "behavior"
 - challenge suspects
- make it hard for sybils to infiltrate the OSN
 - do users care?
 - how can they make better decisions?

DETECTION OF COMPROMISED ACCOUNTS

cost of compromised accounts

- leverage existing trust relationships
- fake account detection not applicable
- cannot be removed easily
- involves costly password-reset process

a recent approach: COMPA

- statistical modeling
 - Extract behavioral profile for accounts
- anomaly detection
 - Compare new messages against observed behavior
- identify campaigns: similar messages & similar new behavior

M. Egele, G. Stringhini, C. Kruegel, and G. Vigna, "COMPA: Detecting Compromised Accounts on Social Networks" in Symposium on NDSS, 2013.

COMPA: example

July 4th 2011, @foxnewspolitics

BREAKING NEWS: President @BarackObama assassinated, 2 gunshot wounds have proved too much. It's a sad 4th for #america. #obamadead RIP

Anomaly scores

- Time: 1.00 (1:24am EST, usually 8-10am EST)
- Source: 0.94 (Web, commonly using TweetDeck) Hashtag: 0.88
- Domain: 0.26
- Mention: 0.67
- Lang: 0.00

COMPA evaluation

Twitter

- Text similarity:
 - 374,920 groups identiKied
 - 9,362 compromised (343,229 accounts)
 - FP: 377 groups (4%), 12,382 accounts (3.6%)
- Landing page similarity:
 - 14,548 groups identiKied
 - 1,236 compromised (54,907 accounts)
 - FP: 72 groups (5.8%), 2,141 accounts (3.8%)

Facebook:

- 48,586 groups identiKied
- 671 compromised (11,499 accounts)
- FP: 22 groups (3.3%), 412 accounts (3.6%)

summary

- why OSNs?
- rewards and challenges of research in OSN
- current research directions
 - de-anonymization
 - privacy (game)
 - Sybil detection/resistance
 - detecting compromised accounts





a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA

Laboratory for Education and Research in Secure Systems Engineering (LERSSE)

Department of Electrical & Computer Engineering

Laboratory for Education and Research in Secure Systems Engineering



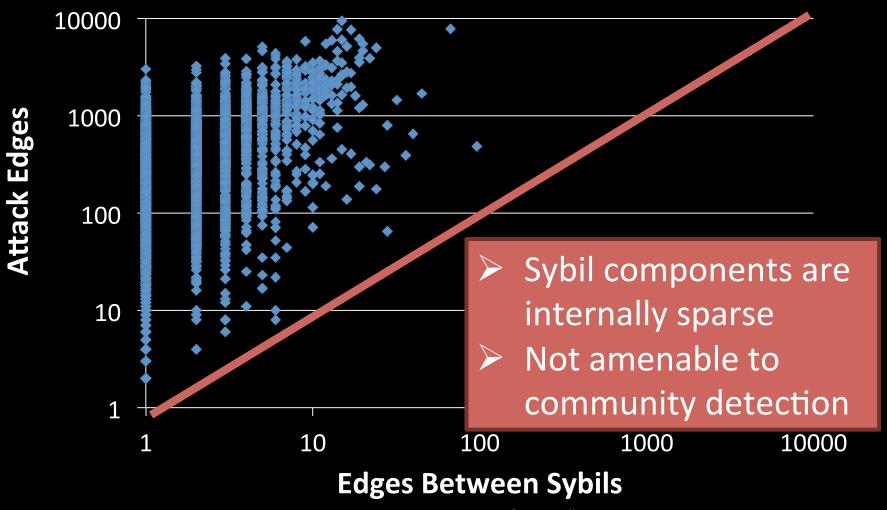
BACK UP SLIDES

OSN size, popularity, and age matter

larger, more popular, and more mature sites

- better privacy protection
- longer privacy policies

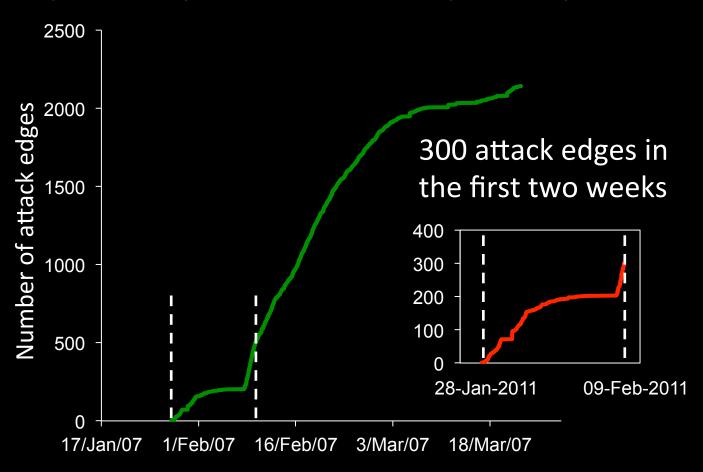
can Sybil components be detected?



Zhi Yang, Christo Wilson, Xiao Wang, Tingting Gao, Ben Y. Zhao, and Yafei Dai, "Uncovering social network sybils in the wild," In Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference (IMC '11).

establishing attack edges takes time

Real-world Sybil activity in Facebook (100 Sybils, fully connected)



Boshmaf et al. Graph-based Sybil detection in social and information systems. To appear in the Proceedings of IEEE/ACM ASONAM, Niagara Falls, ON, Canada (August 2013).



SybilTrack Incremental GSD Algorithm

from graph statistics to graph dynamics

