

Characterizing Privacy in Online Social Networks

Balachander Krishnamurthy, AT&T Labs – Research

Craig E. Wills, Worcester Polytechnic Institute

ACM SIGCOMM Workshop on Online Social Networks

Seattle, Washington USA

August 2008

Privacy on OSNs: Does It Matter?

- Privacy is about keeping *wanted* information.
- With Online Social Networks there is more such information—users are encouraged to (and do) share a range of personal identity-related information.
- Concerns depend on the information, who collects, what they do with it—but most users are unaware.
- Difference between OSN gathering data and third-party sites.
- New wrinkle: External applications popular on OSNs, require users to grant access to all information in order to install.

What Did We Study?

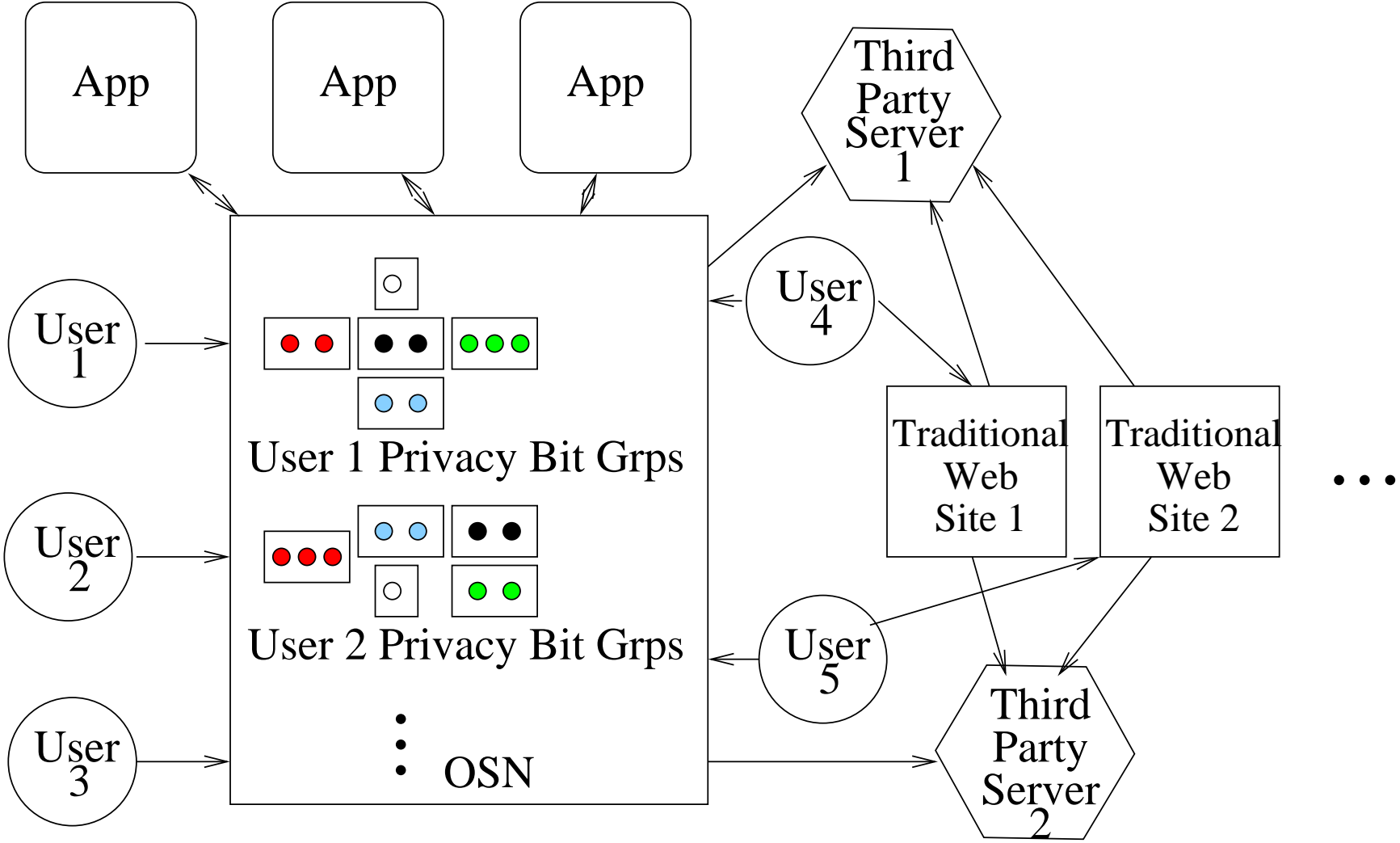
- Notion of privacy bits and groups of bits in OSNs.
- User's privacy controls—what settings are available.
- Default settings and their implications.
- What users do with these settings.
- The role of third party domains in aggregating user-related data and contrast with traditional Web sites.
- Ideas for providing better privacy protection in OSNs.

Privacy Bits

- A bit is a piece of private information
- There are many such bits; e.g., name, age, DOB, a friend, user uploaded content, comments etc.
- Some of the bits can be grouped. E.g.,
 - Thumbnail - name, photo
 - Greater profile - interests, relationship status etc.
 - List of friends
- No way to vouch for the accuracy of these bits or even if the user is a real person.

.

Privacy Information and Potential Leakage



Privacy Groups and Entities

What can be shared:

- Can be ordered L to R by users based on their comfort level
- Groups on the right can be freely shared, on the left not so
- Left may be DOB, right may be name
- They can be stacked vertically if they are equally important

With whom:

OSNs let users to grant privileges to different entities

Typically: user, user's friends, all users

Some OSNs grant privileges to friends of friends or users in a common
“network”

Facebook Privacy Settings

Privacy Bit Group	Self	Friends	Friends of Friends	Friends+ Networks	All
Thumbnail	-	○	○	○	<input type="checkbox"/>
Greater Profile	-	○	○	<input type="checkbox"/>	-
List of Friends	-	○	○	○	<input type="checkbox"/>
User Gen. Content	-	○	○	<input type="checkbox"/>	-
Comments	○	○	○	<input type="checkbox"/>	-

Key: Can be set: ○; Not possible: -; Default:

- Greater profile ('profile' in FB) can be set to be viewable by friends, FoF, or friends and users in same networks (default).
- By default thumbnail and list of friends is viewable by all.
- Networks are an important concept in FB ...
- ... and also a potential source of privacy leakage as it is easy for others to become a member of some networks—users do not know who all can see their information.
- Double vertical line in table is a threshold for user control

MySpace Privacy Settings

Privacy Bit Group	Self	Friends	Friends of Friends	Friends+ Age>18	All
Thumbnail	-	-	-	-	<input type="radio"/>
Greater Profile	-	<input type="radio"/>	-	<input type="radio"/>	<input type="radio"/>
List of Friends	-	<input type="radio"/>	-	<input type="radio"/>	<input type="radio"/>
User Gen. Content	-	<input type="radio"/>	-	<input type="radio"/>	<input type="radio"/>
Comments	-	<input type="radio"/>	-	<input type="radio"/>	<input type="radio"/>

Key: Can be set: ; Not possible: -; Default:

- Coarse-grained settings—all or nothing settings.
- Everyone has access to everything by default.
- Privacy controls for other popular OSNs (Bebo, Digg, Friendster, Hi5, Imeem, LiveJournal, Orkut, Twitter and Xanga) tend to be similar to coarse granularity and all-or-nothing settings of MySpace.

What Do OSN Users Do with These Settings?

'05 CMU study: 1.2% of Facebook users changed thumbnail privacy setting and 0.06% changed profile visibility.

Our two initial studies:

1. MySpace: picked 5K ids at random; 80% did not change default
2. Bebo: 80% did not change default (chose ids among interest groups)

Use of Facebook Privacy Settings

Need to study networks: regional, high school, work

Anyone can join a regional network, others require some nominal ‘proof’ (an organization issued email address).

User can be in one regional network at a time.

506 regional networks (April 2008)

272 in US: cities and their region

234 global networks: cities in Canada/U.K., but countries elsewhere

We picked 20 U.S. and 18 non-U.S by first subdividing each set of networks into four size ranges then choosing specific networks within each range to ensure size and geographic diversity.

Used random network browsing feature of Facebook to obtain users within a network.

Privacy Settings in U.S. Facebook Regional Networks

Regional Network	Users (K)	%View Profile	%View Friends
New York, NY	866	53	78
Chicago, IL	649	54	78
Los Angeles, CA	595	62	82
Atlanta, GA	390	56	82
Dallas/FW, TX	336	63	84
Seattle, WA	210	64	83
Sacramento, CA	99	76	90
Des Moines, IA	83	67	85
Okla City, OK	80	71	87
Greenville, SC	66	72	90
Syracuse, NY	54	75	90
Worcester, MA	45	77	94
Peoria, IL	44	77	93
Boise, ID	36	83	96
Tupelo, MS	29	76	98
La Crosse, WI	25	71	94
Monroe, LA	21	79	98
Ithaca, NY	17	78	95
Abilene, TX	10	82	97
Casper, WY	6	84	99

Strong negative correlation between network size and percentage of users allowing profile and friends to be viewed.

Privacy Settings in Non-U.S. Facebook Regional Networks

Regional Network	Users (K)	%View Profile	%View Friends
London	2486	51	76
Australia	2015	63	83
Turkey	1866	50	76
South Africa	646	65	88
India	633	68	86
Hong Kong	520	59	82
Mexico	448	73	90
Singapore	382	70	88
Greece	241	70	91
Brazil	118	87	96
Edinburgh	98	75	93
South Korea	71	79	88
Jamaica	41	72	91
Iceland	28	84	97
Iran	21	91	97
Algeria	10	92	98
Angola	2	91	98
Nauru	0.2	93	96

Again strong negative correlation between network size and percentage of users allowing profile and friends to be viewed.

Facebook Settings Inference

- Strong negative correlation between network size and user profile visibility strongly. Likewise with viewing friends.
- Same with non-US users—true across cultures.
- Users (apparently) care more about profile info than list of friends.
- Facebook allows further user control of access to some information in a user's profile (e.g. viewing Wall comments). Consequently the privacy of Wall comments is further protected than the View Profile setting.

Use of Third-Party Domains

Performed session of typical actions for each OSN while recording the set of servers contacted for the content of each page.

Executive summary of results:

- Same entities involved as for traditional Web sites (comparison with prior work).
- Users think they are giving information about themselves to their OSN, but others are getting access to what users are doing.

Top Third-Party Domains in OSN Sessions

Third-Party Domain	Online Social Network					
	Fr'ster	Imeem	Bebo	Hi5	MySpace	Xanga
doubleclick.net	X	X	X	X	X	X
2mdn.net	X	X	X	X	X	X
advertising.com	X	X	X	X	X	
atdmt.com	X	X			X	
googlesyndication.com	X		X	X	X	
quantserve.com		X		X		X
adbrite.com	X	X	X			
google-analytics.com	X	X				X
yieldmanager.com			X	X		X

Third-Party Domain	Online Social Network				
	Digg	LiveJ	Facebook	Twitter	Orkut
doubleclick.net	X	X			
2mdn.net	X				
advertising.com			X		
atdmt.com	X	X	X		
googlesyndication.com		X			
quantserve.com	X				
adbrite.com					
google-analytics.com				X	
yieldmanager.com					

What Do We Propose?

- Break privacy into groups
- Let users order them in terms of importance for them and specify how far in the L-R spectrum they are willing to allow access by default
- Let applications and OSNs ask for bare minimum and never more than supremum
- If needed information is within default, access is transparent
- If more is needed, user is asked

Summary

Have examined range of OSN privacy controls, defaults and what users do with these privacy controls—not just a point-in-time study.

Found that controls are generally coarse-grained, provided a limited range of entities to which access can be assigned and allow permissive access by default.

Between 55 and 90% of users in OSNs allow their profile information to be viewable and 80 to 97% of users allow their set of friends to be viewed.

Strong negative correlation between regional network size in Facebook and the use of these privacy settings to limit access.

Much like traditional Web sites, third-party domains track user activity pervasively in OSNs.

Proposed idea to better match what information a user makes available with what is needed by other users and applications.