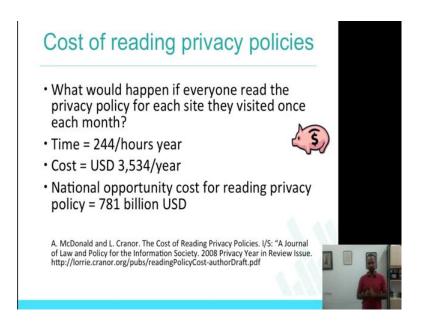**Privacy and Security in Online Social Networks**
Prof. Ponnurangam Kumaraguru ("PK")
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Week - 7.2
Lecture - 23
Link Farming in Online Social Media

Welcome back to the course Privacy and Security in Online Social Media, week number 7 - second section of the week number seven. So I hope you got a chance to look at the content that we made <mark>available</mark> for the week 7.1.
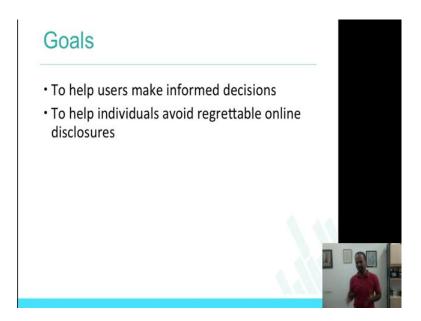
(Refer Side Time: 00:27)

I do not know, how many of you actually read privacy policies. Researchers have shown that people do not read privacy policies. Just let us do some exercise now. So, some of you may have actually done online transactions in the last one week or almost all of you would have done some kind of online transactions, logging to your Facebook, logging to your Twitter, transactions of buying things online or between banks. How many of you have ever read any privacy policy of the websites that you have interacted with in the last week or a month, must you really <mark>low</mark>. And it is not just what I am saying, people have actually studied this exhaustively.

So, here is one piece of work which says, if we were to make people read privacy policies, what it would it cost. The study was done in the US(Refer Time: 01:19) and it is actually part of PhD thesis work, where the question was, what would happen if everyone read the privacy policy for each website they visited once each month. Time taken would be about 244 hours per year, which is basically, that questions more on the lines of economics of what would it cost. It is 244 hours, convert it into money,

, and in total number, if we just look at the US population, national opportunity cost for reading privacy policy would be 781 billion US dollars, which is, if I get every citizen of the US to read the privacy policy for least one month for the websites they have seen, it would cost in some 781 billion dollars. That is a lot of money, and that is a lot of opportunity that is being lost because they are spending time on reading this privacy policy. While they could make the pair, which they would (Refer Time: 02:14) use to make the decision.
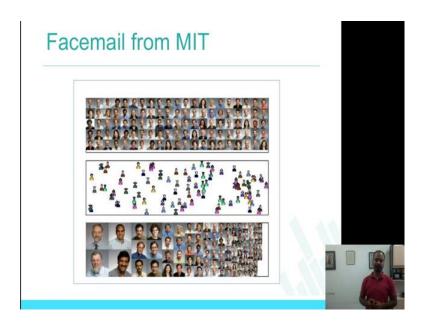
(Refer Side Time: 02:18)

So, keeping that context in mind, which is reading the privacy policies, researchers started asking some questions, and broadly also, there is this whole area where researches are working on, and technologies is being built to help users make informed decisions, which is how can I help users to make informed decisions with the

information that is presented to them, with the information that they can actually use from these services.

The specific goal is to help individuals avoid regrettable online disclosures, which is, can we actually build technology, can we actually build something for the users to use, so that they can actually be happy about the content that that they are posting or avoid being regrettable for the information that they are disclosing online.

(Refer Side Time: 03:10)

So here is one technology that was built at (Refer Time: 03:14) MIT, meaning even now when we are using this discussion forum for this class, when you type in the discussion forum, add the course name, you will actually find out in this plot, when if you use this technology, you do not really know who are the people who are getting this email, right, (Refer Time: 03:36), because the groups are set up differently, people have been signed up into the groups, you really don't get to know who you are interacting with (Refer Time: 03:44). For example, in a mailing list in your company and mailing list in your college or mailing list that you maintain for yourself with your friends, so all of this you do not really get to see who is getting these emails (Refer Time: 03:57).

So, what these researchers at MIT did was, they are actually said, okay, whenever the

email is going to go, because this would, this could, be a problem also, right, because you do not know who is getting the email. So, it could be a problem the information that you are sharing could actually go to people whom you do not want them to see this content.

So what they said was, okay, if there is a email address the state called (Refer Time: 04:19) psosmnptel2016 at abc dot com, they would actually show you the profile pictures of the people who are going to get this email, so that is a way by which to show that how many people are actually going to get the email, that is the information on the slide also, their profile pictures. And it would also show you who are the people who are getting the, because of the profile you get to know more about the people who are getting to see this email.

So, this just helps users to make a decision on sending this email, because let us take if you were to send an email, and if you wrongly typed the mailing list address, it could actually end up going to a wrong mailing list. You wanted to send it only to 10 people, whereas the email list that you send is actually going to 100 people, so this can actually help avoid (Refer Time: 05:10).

In another version of it that they built also where they were actually showing you the bottom one, when they were actually showing you the profile picture of the person who you have been interacted more slightly bigger than ot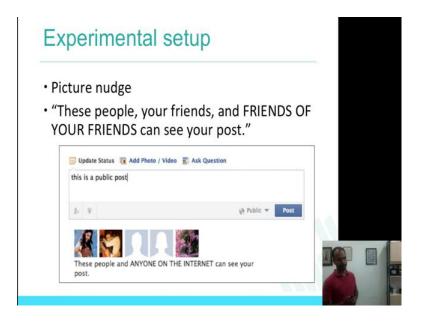hers ones. This is also helping you to make a judgement on who is getting the email, who were the people who are getting these e mails and who you have interacted with more frequently than others. So this is just an another example of helping users to make informed decision, whether you want to send the email or not.

(Refer Side Time: 05:48)

So, keeping that in mind, which is to help users make this decision - informed decision - here is a piece of work that was done to look at the context of just Facebook. So in this experiments that they did, what they did was they built Chrome browser extension, which would actually show you, you go to facebook.com, when you are going to do a post it is actually going to nudge you with some information. They have different set of sub nudges, I will walk through what they did.

And we will also see how effective it was, which was effective, which was not effective. Because today, you could actually, meaning, I am sure some of you have experienced that you did some post and which you did not want somebody to, somebody in your friends network to see, and they got to see the post, and there have been incidences in the past also. If you remembered even in the week 1 or week 2, I showed you some examples about MI6 chiefs (Refer Time: 06:44) while posting some content that went public, where it was not intended to go public.

So, these kind of incidents have actually made people to think about building technologies that will help them, help users make that decisions. So here is the first idea called picture nudge. What it does is if you wanted to do a post, when you are doing a post it is actually going to stop you and say that these people, which is the profile picture connected to the face mail, it is the same thing as in the email mailing list. These people your friends and friends of your friends can see your post, it is going to stop you and tell
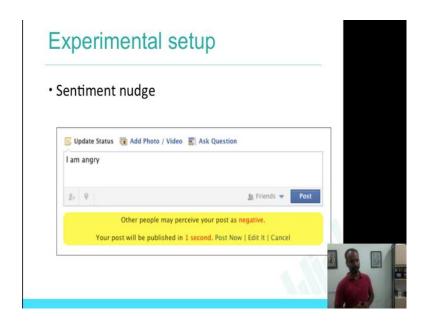
you this information. Again in the bottom screenshot which showing you these people and anyone on the internet can see your post. So, this actually helps you to make a decision on whether you want to actually do this post or not.

(Refer Side Time: 07:38)

Here is a second experimental set up which is timer nudge. Here it is not showing you the profile pictures. But it showing you that you have ten seconds to cancel your post. This information just lets you to say that essentially, you are doing this post, you really want to do this post, wait for ten seconds, if you want to change your mind, do it now, and then do the update. That is the timer update.

(Refer Side Time: 08:10)

Experimental setup

- Sentiment nudge

Here is next one which is sentiment nudge. Again there have been incidences all around the world where people have actually posted the things on Facebook and it is actually backfired on them. Backfired in terms of actually very negative effects also for the content that they have posted.

So, to avoid such things here is a setup where it shows you and in this case I am angry, so it shows you that it is negative; other people can perceive your post just negative. So, negative sentiment is attached to the post, so please be careful, do you really want post it, and things like that, so that is the sentiment nudge. Let us look at all the three - picture nudge with pictures, timer nudge with actually ten seconds time, and sentiment nudge which gives you the sentiment of the post that you are making.

(Refer Side Time: 09:04)

So methodology of the study that they did was Chrome browser as they said, they did exit surveys, which is when people completed the study, they asked them some questions and also asked them questions in terms of both quantitative and also interviews with the participants. So, IRB is institutional review board, which is basically to say that if you are interacting or collecting data from users, it will be human subjects, you really want to make sure that things do not go wrong when they are actually doing the study, and they should not feel offended and things like that. And that is why IRB approval is necessary when you are interacting with human subjects.
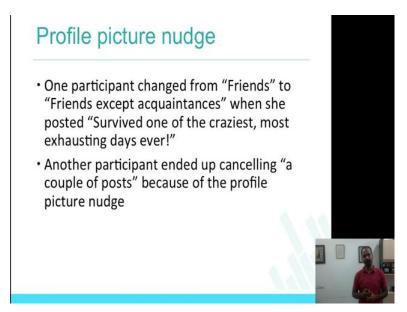
And of course, users were recruited in multiple ways. (Refer Time: 09:52) Of course, putting flyers all around the place sending out e mails, Craigslist, all of that. They got 21 participants who completed the field study, because it is the Chrome browser plug-in, they could actually use it at home or wherever, and 13 people participated in the interviews, which is the exit survey.

(Refer Side Time: 10:12)

Analysis metrics

- Number of changes in inline privacy settings
- Number of cancelled or edited posts
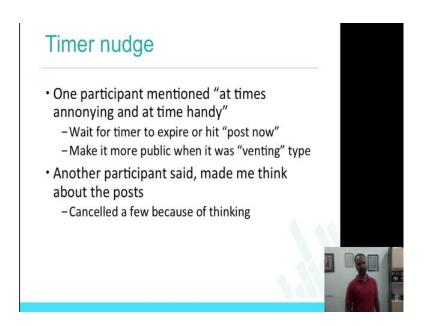- Posts frequency
- Topic sensitivity

So, in the metrics that they are actually used to analyze what is going on within this context of providing nudges of these things. Number of changes in inline privacy settings, number of canceled or edited posts, post frequency and topic sensitivity, essentially they were trying to understand by giving these nudges are people changing the behavior. And if they change the behavior what are they changing, so that is the concept that they were actually trying to study, that is the metrics that they were trying to study.

(Refer Side Time: 10:43)

Profile picture nudge

- One participant changed from "Friends" to "Friends except acquaintances" when she posted "Survived one of the craziest, most exhausting days ever!"
- Another participant ended up cancelling "a couple of posts" because of the profile picture nudge

Profile picture nudge. So the first one, I think I have one slide for a nudge to tell you what happened in this study. One participant changed from friends to which is probably all of them, friends except acquaintances, when she posted survived one of the craziest, most exhausting days ever. So essentially people are changing the group in which they are sharing the content depending on the information that the nudge is actually providing them. Another participant ended up canceling a couple of posts, because of the profile picture nudge basically saying that oh it is going to actually lot more people than what I thought, what I think, so let me not do (Refer Time: 11:22) the post.

(Refer Side Time: 11:24)

Timer nudge, one participant actually said at times annoying, and at times handy. Because I am pretty sure, right, because if you are doing, let's take, 10 posts a day or, like, 5 post a day also, it is going to stop you for every post ten seconds and then only to go. Waiting for a timer to expire or hit the 'post now'. So it is essentially a feature that could be provided. Make it more public, when it was venting type, right, make it more, so, it can be more suggestive. Another participant said, made be think about the posts, which is the same, which is the behavior that is actually important for these nudges to actually make within the users(Refer Time: 12:05). Changing the user behavior, it is a big need while using this technology.

(Refer Side Time: 12:15)

Sentiment nudge, nudge was missing the context, of course, right, because I think the whole context of, in what context I am doing the post is actually very important to find out sentiment. And even human beings, it is actually hard to get the sentiments, so, the nudge was, the tool was actually making, browser plug-in was actually making errors while the calculating or finding out what these sentiments are.

Many participants canceled the posts, because, I think it is because if the posts are negative, and many people are going to be actually offended by the post within your network or in public, it is actually going to be bad for you. So, people are actually canceled the post. And the post frequency also reduces; they were actually doing 13 posts, it went down to 7.

(Refer Side Time: 13:07)

Conclusion

- Interventions help users make better decision
- More work is needed in order to understand which type of nudge works in which context

So, essentially all of this is helpful in making the, helping the users to make better decisions. And particularly when it comes to posting something that others are going to be offended, posting something that people are going to, people whom you do not intend to actually share the information, all of this is actually helpful, these technologies are helpful, for the users to make a better decision.

And of course, more work is needed to understand which type of nudges work in which context, because the contexts could be very different - I am just doing you a quick update, saying, I am, for example, I was actually doing a lot of updates, yesterday, about the convocation at IIIT-Delhi with the hash tag IIIT Convo 5. And if I wait, if it was (Refer Time: 13:55) going to stop me, and probably I did like thirty, forty posts on Twitter. But every time if it is going to stop me for ten seconds, that is not going to be good, so, and also, in terms of the, in terms of the sentiment that it is showing, everything has to be done slightly better.

So, I think more work is needed, I am sure there are people in the class who are interested in taking some of this, it may be interesting projects to work on. With that I will stop the 7.2 part of the week. And I will continue on something which is in the context of phishing; and phishing in the context of social media in the next part of this week also.