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

Week - 7.3
Lecture - 24
Semantic attacks: Spear phishing

Welcome back to week 7 and this is the third part of the week 7. In this class, in this section what will see is, we will see about Phishing Attacks in online social networks.

(Refer Slide Time: 00:18)

So, this is a slide from an MIT PhD thesis which actually looked at what a semantic attack is? Semantic attacks are attacks that happens where humans are targeted. So, for example, Bruce Schneier who is supposed to be a expert in security classified the different types of attacks that could happen as physical, syntactic and semantic, but physical attacks are the were happening like 15-20 years before where the attackers would actually get physical access to the machine.

Whereas in syntactic attacks are the attacks that were happening around the programs, around the systems that are built, which is more like the denial of service attacks, buffer overflow attacks and attacks like that, but the semantic attacks are attacks which target

the way we as humans assign meaning to the content which is that what do we because the specific attack that we will be talking about is phishing.

For example, if you get an email from pk at iiitd dot ac dot in now, talking about NPTEL course and which has a link saying please give your user name and a password to see the content here most likely that you're gonna actually click the link and give the information which may be a phishing link also. So, that is started the way you actually think you are seeing an e mail that is coming from legitimately pk at iiitd, and the system thinks that you are actually going to this free website forum psosm on NPTEL dot come slash login dot html, but actually it is a phishing website. it is targeted phishing website.

So, system and mental model, what this is in this PhD thesis they actually nicely put it that semantic barrier, which is the difference between what system thinks we are doing and what you think that system is doing will actually be called semantic barrier in the larger the barrier is it is because actually difficult to not fall for such types of products. So, if you look at the mental model it says, who is the other party what is the meaning of the message. So, the example that I said also what is the meaning of the email we got what is the meaning of the who is sending the message and information is all mental model of the user, but a system model who is the remote machine where booked website I am going to access and information like that.

So, user model or the mental model which what users think that is happening, system model which is what, the system thinks that the user are doing, the barrier between them the difference between them is actually called semantic barrier. and the larger the barrier is, its actually hard to actually fix the problem. So, this is what we will use this is what we will actually talk about mostly in the section called phishing.
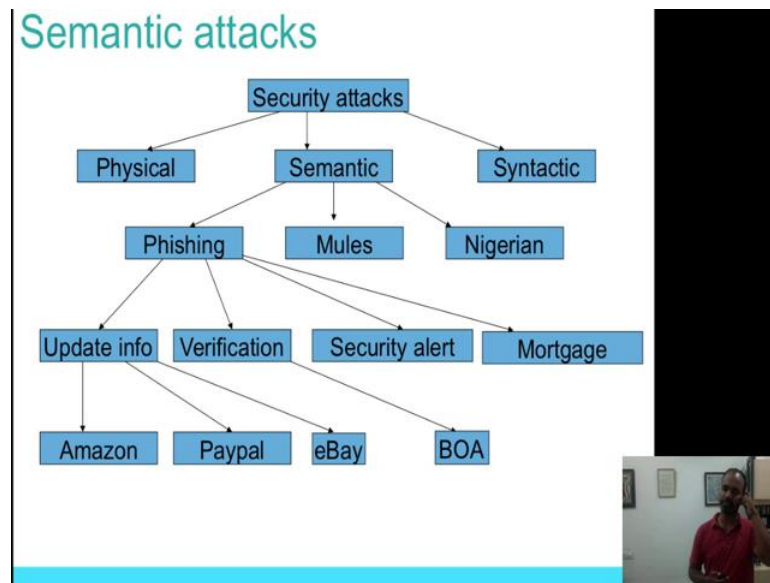
Here is the broad category of semantic attacks: Security attacks - physical, semantic and syntactic which is what Bruce Schneier did and if you look at Semantic attacks, you can actually go through multipe categories Phishing, Mules, Nigerian, 4 1 scams and attacks like that, and in phishing also there are multiple categories – update your information, banks and in your ICICI banks sending you a message saying that, please update your information within next 24 hours or your account would be closed. Verification, saying that, we want to verify whether it is really you, please click and verify. Security alert, Microsoft is updating the latest version of MacOS, there is an update.

Here is the link, please go and update. Mortgage information, meaning your mortgage, the due is coming closer, please click this link and do something. All of these kinds of categories of attacks are called phishing attacks and almost all companies today probably are undergoing, are part of, or being victims of this attack of phishing. Even academic institutes probably are victims of phishing attacks.

Here is a simple example, which is an email that the 3 parts of the email which is actually makes the legitimate email and the difference between the legitimate email and the phishing email, which is the subject line, subject line and urgency in the message on there are there is the line. These are the three things that happen that is a part of the phishing email which at least one wants to keep attention on. Subject eBay urgent notification from billing department. We regret to inform you that your eBay account could be suspended if you do not update your account information and then there is a link there and then when we click on the link it takes you to a website called kusi dot org.

(Refer Slide Time: 05:15)

Which supposedly should be taking you to eBay sign in page, so that is the sharing that is a very classical phishing attack when there are multiple ways of a changing these kind of phishing attacks.

(Refer Slide Time: 05:34)
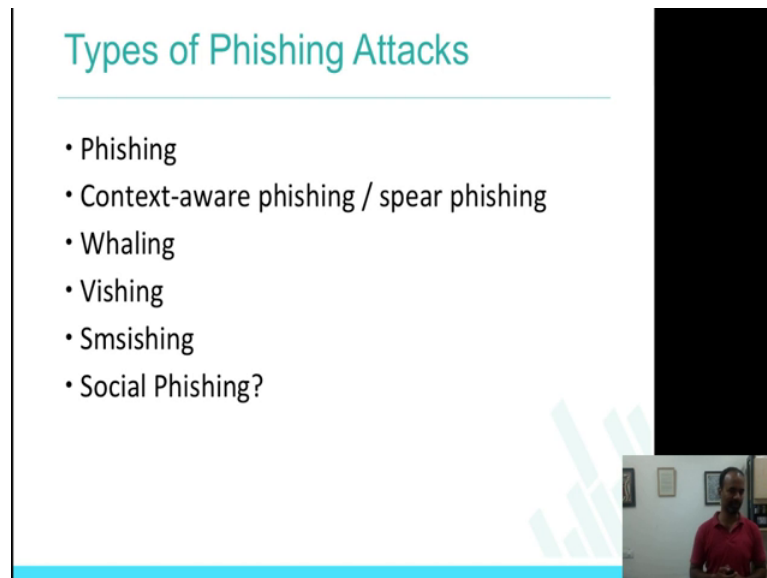
Here is some cost again, economics about phishing, some costs that is relevant to the topic of phishing. Costs of hundred thousand employees organization, which is the, if the phishing attack happens what would be the cost to contain the malware, the cost to contain a malware, the cost of malware not contained. So, if you look at the cost its actually pretty high in terms of actually even the phishing attack. Total extrapolated cost is 3 million 76; 3 million plus dollars, all right. So, it needs a lot of money that is spent every year, FTC and many other organizations in US actually try to course about phishing and there is an organization called anti-phishing working group which actually specifically works on the problem of phishing and how to actually reduce it.

(Refer Slide Time: 06:28)

So, here are some kinds of Phishing Attacks I think we probably briefly mention this in the past also. So, I'll go over quickly, phishing which is a classical one that I showed you, Context-aware phishing the email that I talked about sending in to the students taking this course, Whaling is an attack which is sent to the chief executive officers of the company, Vishing is over the phone, Smsishing is over the SMS. So, what is social phishing? That is what the topic that we have been discussing for the rest of this week.

Social Phishing; does anybody know what social phishing is?

Social phishing is nothing but looking at the information from the social context then using that to actually phish, it is not about finding whether you are taking a course that could be many other information that I defined on from a Facebook page, from the facebook account, things that you've done and things like that. So, the topics that we have seen until now are using older data we saw Latanya Sweeney's work using medical health data, again Latanya Sweeney's work, using pictures from FB voter data.

We also saw the work that was down in collecting pictures from the university campus collecting this information and making some judgments about the user. Finding the people who they see in the campus whether they will get the right profile from the Facebook. So, those are the topics that we saw, but we never saw about what social phishing is.

So, here is a goal the goal is to see how phishing attacks can be performed by collecting personal information from social networks right. So, it is not about sending into the CEO's, it is not about sending to the students of this class. It is about actually can I collect some information about you from your social network behavior and use it against you, how easily or effectively can phisher use this information. Again there is a very classical work that was done some years back. So, it'll be nice to actually know how they did it some years back. I am pretty sure these studies can be done again to see how it goes and this study was done in the US.

(Refer Slide Time: 08:41)
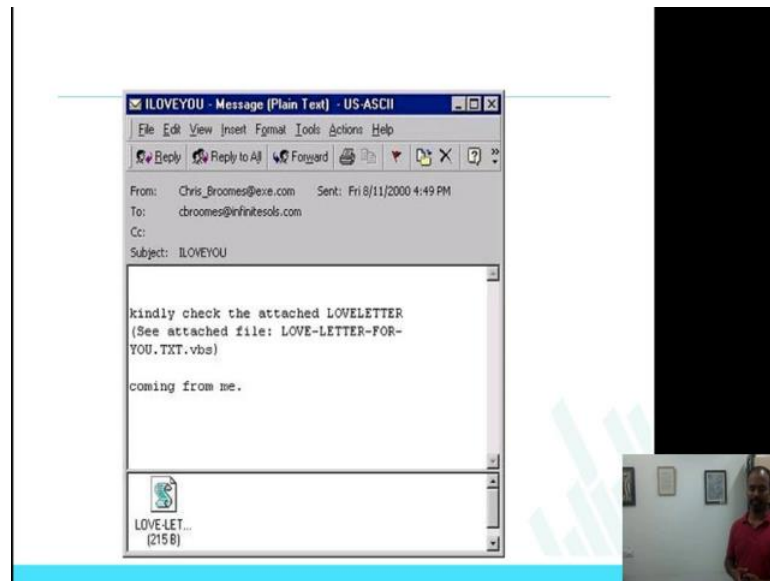
Here are some examples. So, I love you virus, some of you may know this. Kindly check the attached love letter, see attached files, this is an email that comes, coming from me right. It was one of the first virus that was actually spread.

(Refer Slide Time: 09:02)

So, what they did was they collected publicly available personal information using simple tools you could actually collect now information from Facebook. So, we will actually have some tutorials also about NLTK, how to use it how to analyze this text that are coming from these post, all right. So, you could collect this information and find out what is the date of birth mentioned there. This was done in Indiana university. I was referring to Indiana university. Coerrelated this data with Indiana University's address book.

Which is they collected all the posts done by students of Indiana university and then they launched the study in April 2005, they launched for the age group between 18 and 24 which is the student population in campus most of the times.

(Refer Slide Time: 09:53)
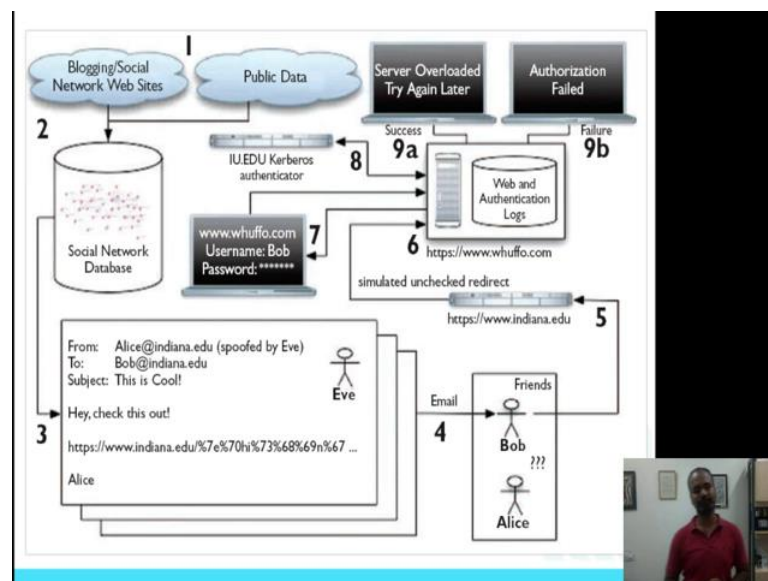
So, here is the slide which actually I think the next slide we have also, go through this in a text form, but here is the slide that actually walks you through in terms of what the procedure that they follow right. First they actually look at public data, blogging, social networking sites, they collected the data which is stored into the social database, social network database, they use this data to create an email which is - From Alice at indiana dot edu, To subject Bob at indiana dot edu. This is cool, hey check this out with the URL there, right.

So, from there the information is sent as an email, bob to friends at Indian University, and when the user clicks on the link. It goes to the Indiana University's website and it actually checks authentication web and the authentication logs, it tracks and takes into a user name and a password page, then when they give user name and password it checks the user name and the password whether that is appropriate which is checked.

Controls authenticator and comes back and server overloaded try again message is sent, authentication failed. So, those are two outcomes of the whole process which is success, server overloaded try again and an authorization failed right. That is a process that they followed in terms of finding out information from social networks, sending out this emails and getting some uses to get to that page. Many people that have done this study after that, even if you go look out my own work in 2007, 8 and 9, I have done similar kind of sending out phishing emails and seeing how people behave.

(Refer Slide Time: 11:35)
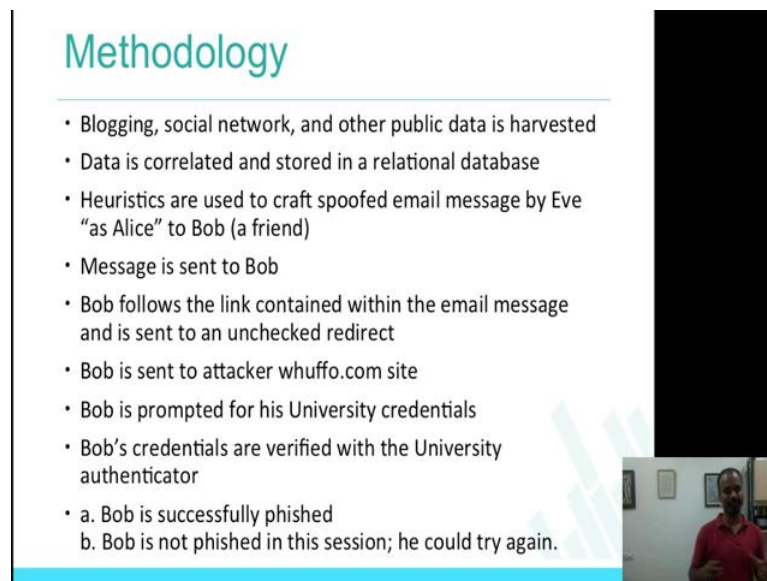
So, two things that they did, of course it was experimental reseatch. So, they were trying to compare how the email from Indian university email id, but from an unknown person that is a control goal, if I get an email from Indiana email id which in my case I get an email from somebody who is from IIIT, Delhi with the iiit email.

But I do not know the person because that's something I can actually get from the social context, experimental group is from a friend in Indiana university itself, which is if I have already connected to the friends in some way you saw the Facebook posts and following this person in twitter, I mention them in the post on twitter, so all that it actually helps to find out that is the experimental setup.

(Refer Slide Time: 12:24)

Methodology

- Blogging, social network, and other public data is harvested
- Data is correlated and stored in a relational database
- Heuristics are used to craft spoofed email message by Eve "as Alice" to Bob (a friend)
- Message is sent to Bob
- Bob follows the link contained within the email message and is sent to an unchecked redirect
- Bob is sent to attacker whuffo.com site
- Bob is prompted for his University credentials
- Bob's credentials are verified with the University authenticator
- a. Bob is successfully phished
  b. Bob is not phished in this session; he could try again.

So, the same methodology, the chart that was there here is the verbose of it, blogging social network and other public data is harvested, data is correlated and stored in a relational database, heuristics are used to craft the spoofed email messages, message is sent to Bob, Bob follows the link contained within the email and is sent an unshared redirect, bob is sent to an attacker whuffo dot com.

Bob has prompted for his university credentials, bobs credentials are verified with the university's authenticator and bob is successfully phished, bob is not phished in this session, he could try again all right. That is the verbose of the architecture that was shown or the experimental methodology that was shown in the slide before.

(Refer Slide Time: 13:13)

Continuing on the analysis of data that researchers collected in terms of social phishing here are the results. So, in this table what we are seeing is control condition and social condition experimental condition, as in the rows columns being successful targeted percentage and confidence intervals. Successful meaning how many people got those emails who actually fell for the it, targeted the number of peoples who were actually sent this email to percentage of course, is the ratio of targeted versus successful. So, let us look at some results.

(Refer Slide Time: 13:49)

So clearly control group is, high which is that 16 percent of the participants falling for these kind of emails its very, very high in general it is not a bad level when in I think it is hard to believe that 16 percent of the participants actually fell for this kind of these email, but the advantage here or the context to keep in mind for the data is that sender email was from Indiana university itself, I think that is the reason why this percentage is very high.

For example, if you get an email from pk at iiitd dot ac dot in versus if you get an email from pk at lets takes some abc dot com, the higher chance of you clicking and going through what a email saying, asking you to do would be pk at iiitd dot com or pk at iiitd dot ac dot in, is very high and of course, 72 percent of participants in the social condition clicking the email and doing whatever is asked in the email is actually pretty consistent with other studies that have been studies where they have shown that this percentage is even higher than 72 percent.

More results which is to see the success rate of how people, authenticator to this website. So, here are some interesting results again. Seventy percent of authentications. So, what does this graph show, this graph has x axis being the time, date, the dates which is 6 pm, 6 am, 12 noon, 6 pm as in the x axis, y axis to be the percentage of people who actually

clicked for authentication, so green is showing you cumulative authentications, red line showing you authentications per hour and then blue line is visits per hour. So, you essentially what does it mean, blue line is showing you that number of people who went to this website, red line is showing the number of people who actually authenticated which means it'll always be below.

This clearly shows that 70 percent of authentications in the first 12 hours. So, if you look at the first part of graph, 70 percent of authentications which is the red line which is actually in the first 12 hours itself. The problem is that people fall for these kind of attacks immediately when they get these emails, right because there is a sense of urgency, there is a sense of completion, completing it immediately and that level of urgency is put in this email. If you remember the example that I showed you from eBay website, an email that they sent has a subject line also has urgent notification, urgent verification, but this actually puts a challenge on solving the problem.

Phishing, which is takedown has to be successful, which is if the websites are taken down as early as possible as soon as possible, then there is a high chance of this several users who were going to this website can be actually stopped. If the websites are not taken down unfortunately these users just actually end up actually going to the fake website and giving away other personal information right. Again success rate, how people react to these emails what level of authentication, what is the percentage of people who are actually giving their account details, is what they show in this graph.

(Refer Slide Time: 17:17)
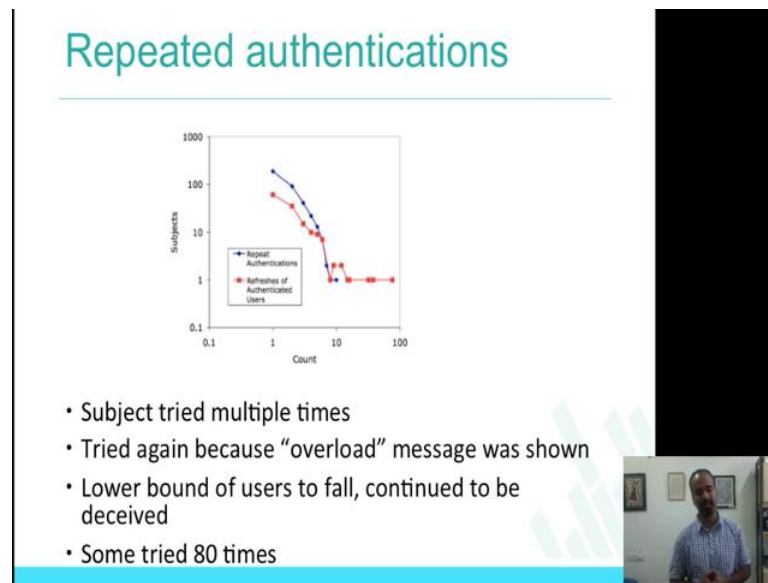
So, here is another interesting analysis that this research actually shows which is that subject, subjects actually tried, participants tried multiple times to actually authenticate which is the blue line is actually showing you repeated authentication and the red line is showing the refreshes of authenticated users, which is they trying to refresh and see whether they are able to log in to system. If you remember the architecture that final output is two. So, let me go back, final outcome of the study is at two levels, where this authentication failed, server overloaded, so they try again. So, when user sees this they feel something is wrong with the system. Let me just refresh it and let me just try it again, that is what is happening in here.

So, tried again because overload message was shown that lot of people who actually tried because the overload message was shown. So, this is basically showing you the lower bound of users to fall and continued to be deceived and if you we look the blue line which is people are actually authenticating to this website repeatedly, even it is actually showing you that authentication error, some people actually seem to tried to 80 times.

So, the x axis is here showing you the count which is the log scale, y axis showing you the number of subjects you can clearly see that about 80 percent of the 80, some people even tried it for 80 times to get into the website, and this is not the only study, which is

showing this, this is probably the classical study, one of the first studies which showed this, but later there have been many studies who showed that such kind of repeated authentications happen with the users.

(Refer Slide Time: 19:02)

Gender

	To Male	To Female	To Any
From Male	53%	78%	68%
From Female	68%	76%	73%
From Any	65%	77%	72%

- 18,294 Ms and 19,527 Fs
- Overall F more victims
- More successful if it came from opposite gender

Here is the ratio, here is the analysis of the gender, because they had the Indiana university's, university student details they could actually find out male versus female, the gender details of the participants. So, this table actually shows you on the rows, it shows you from male, from female, from any one, to male, to female and to anyone. This basically says that if the email is coming from, again if you remember the study was set up, they collected the data, they have crafted the email, and while they are crafted the email they were actually doing all these experiments to see, if I send it from male what happens to when the email goes to a male versus male to female, all right.

So, this shows that overall female were more victims which is you can see on the third column, which is to female being much higher than to male, it does not matter where the email is coming from, female seems to be more vulnerable to these kind of attacks. 18,294 males and 19,527 females were actually being part of the study, more successful if it came from the opposite gender. You can clearly see that from male to female which is the row 2 and then the column 3, 78 percent and from female to male which is 68

percent. So, this number which is 68 is the highest in the column of the to male, 78 which is the highest in the column of to female, which basically shows that if a male gets an email from female and the female gets the email from a male, it is actually high.

The percentage of chance of actually authenticating and giving away the information is much higher, if the email came from the opposite gender, that is a very interesting conclusion that to show that phishing attacks, or vulnerable phishing attacks are successful, but is also more successful if the emails come from the opposite gender, and sure these results can be repeated even in non email context.

(Refer Slide Time: 21:18)

· Younger targets more vulnerable

Again given that they had a lot of demographics data here are some things about age group, things about the departments that they were part of. You can clearly see here that the younger targets are more vulnerable, which is the younger and the participants are the more vulnerable that they are to, for authenticating to the study. You can see that freshman, the difference here is that the orange and the blue, the orange is showing you the social phishing which is the experimental setup; the blue is showing you the control condition, while you can see the difference between the freshman, difference between the social and the control is the highest in terms of freshman.

And as you go up it keeps reducing, so from or junior and senior in sophomore, it looks like little high, but if you put the sophomore and freshman together it basically says that the younger the people the more vulnerable they are to these kind of attacks.
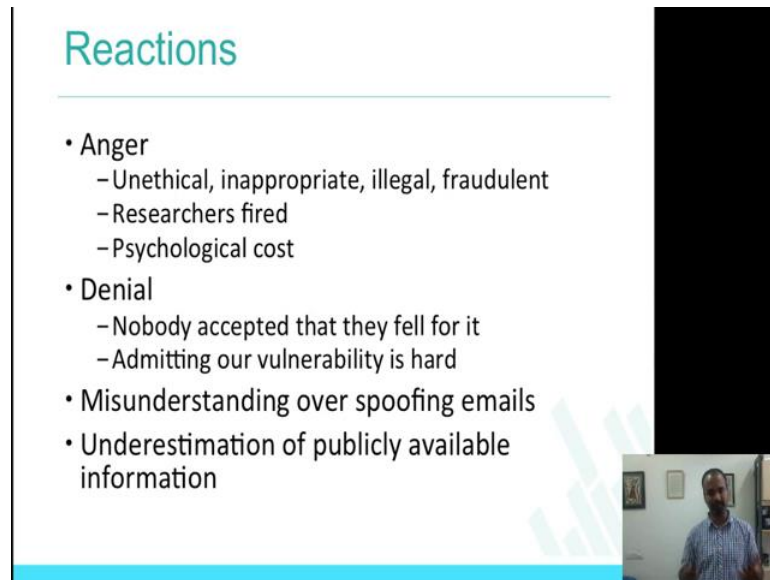
(Refer Slide Time: 22:18)
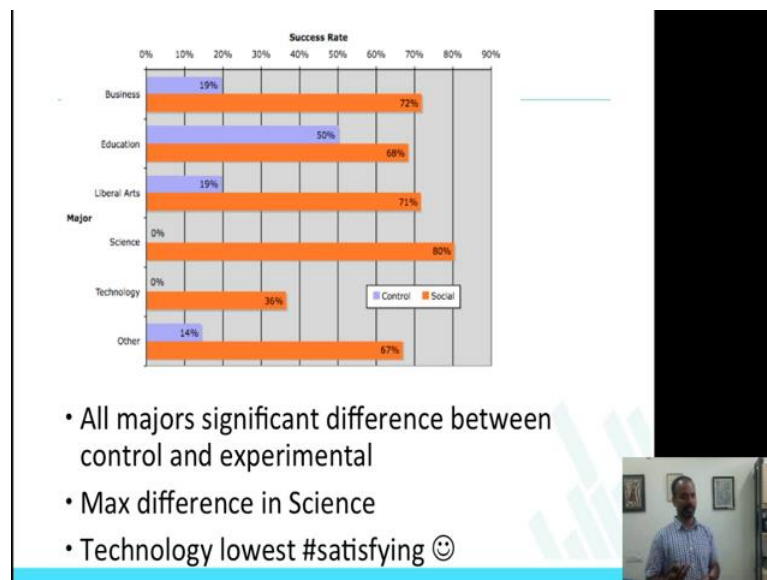
Similarly, they also had which department the participants came from. So, they were able to actually create a graph which looks like this, again the same color of color scheme, which is orange being the social condition and blue being the control condition, which basically shows that all majors significant difference between control and experimental, which is any department of the campus it does not matter, the difference between social and controls is very high which is social people fall more compared to the control condition.

It also showed that the science department had the maximum difference, if we look at science 80 percent is for the social and 0 percent is for the control condition. So, which shows that the science department had the maximum difference between the social and the control. And it was also evident that the technology has the smallest, which is people who study technology that have probably are less vulnerable to these kind of attacks which is about 36 percent here, difference between the social and the control condition right. So, this is way by which research is actually found, which kind of department and

the students going to which kind of departments are actually vulnerable to these phishing attacks.
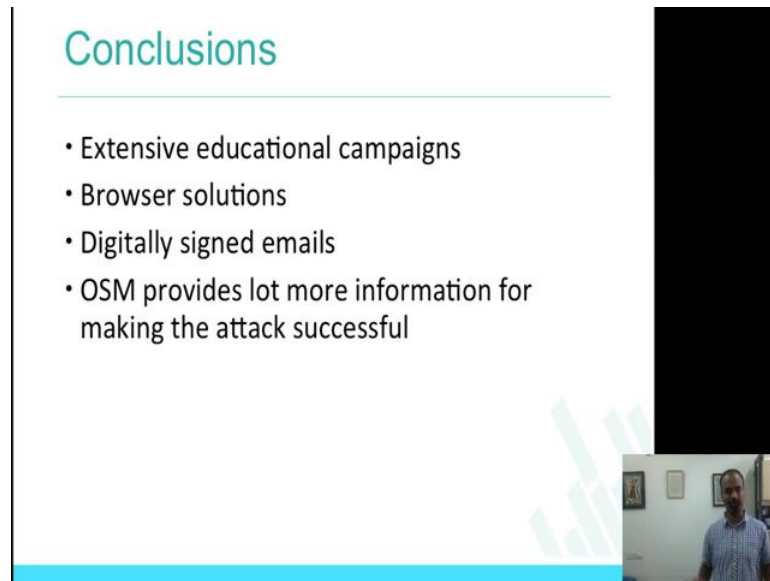
(Refer Slide Time: 23:38)

In general, this study got a lot of negative reactions from the participants which is like it was unethical, inappropriate, illegal and it was also fraudulent, researchers fired researchers were fired, psychological participants claimed that there were psychological cost, because I think they were under pressure, they did not know about the study happening and things along that, and interestingly there were people who wrote blogs, people who wrote reactions about the study and they said that they were not part of the study and they did not fall for these attacks with somebody else fell for, which also shows that admitting that I am vulnerable is actually is also hard; I think that is a misunderstanding over spoofing emails, underestimation of publicly available information.

So, participants did not, meaning generally also you and I will not perceive how bad the publicly available information about you can be used against you, since this was one of the first studies these reactions were actually interesting, but there are people who have done the studies after this which were again you studied how people fall for phishing emails.

(Refer Slide Time: 24:52)

Essentially, what does the results show is that extensive education campaigns is necessary, browser solutions of course, take down has to be much faster, digitally signed emails have to become more prevalent and of course, online social media provides lot more information for making these attacks more successful, all right. So, people should stop sharing a lot more personal information on social networks, digitally signed emails should become more prevalent, browser solutions should be built.

To say that this website is actually, this email is actually phishing and this website is actually malicious and this website is a fraudulent website, and of course all of this can come to education campaigns.

(Refer Slide Time: 25:39)

Some reference is for this research that I discussed. With that I stop actually the week 7; we will actually look at some more exciting topics in week 8.