Inadvertent Data Disclosure on Peer-to-Peer Networks
Transcript

Part 1: Understanding the Threat

Julia Allen: Welcome to CERT's podcast series: Security for Business Leaders. The CERT Program is part of the Software Engineering Institute, a federally funded research and development center at Carnegie Mellon University in Pittsburgh, Pennsylvania. You can find out more about us at cert.org.

Show notes for today's conversation are available at the podcast website.

My name is Julia Allen. I'm a senior researcher at CERT, working on security governance and executive outreach. Today I'm pleased to introduce M. Eric Johnson, Director, and Scott Dynes, Senior Research Fellow, both with the Center for Digital Strategies within Dartmouth's College Tuck School of Business. Today we'll be discussing a growing class of information security breaches caused by inadvertent disclosure of sensitive business information. So Eric and Scott welcome. Thanks for your time today.

Eric Johnson: Thank you.

Scott Dynes: Thanks for having us.

Julia Allen: In a recent paper that you both published you discuss the fact that while intruders do target poorly secured networks, many security breaches today are due to inadvertent or unintentional disclosure. Eric, could you give us some examples of this?

Eric Johnson: Sure. Well so much of the press and media has been focused on what we call hacks, but many times when you look more closely these are not technical hacks. They are much closer to what we call inadvertent disclosures. And these fall in lots of different categories, everything from lost laptops to mis-posted entries on the web, to losses in peer-to-peer file sharing networks. And in each case the result's the same. Information is leaked out of organizations and many times to the embarrassment or financial loss of some of the stakeholders.

Julia Allen: So you mentioned this, Eric briefly, that a growing area that we've seen and that you've both researched or written about quite broadly for inadvertent disclosure occurs in peer-to-peer file sharing networks, and that's where we're pretty much going to focus our conversation today. So Scott what is this and why is it growing?

Scott Dynes: Well peer-to-peer networks, examples include things like Napster, previously LimeWire, eDonkey — essentially these are networks where people typically share music and you've read a lot about the music companies getting upset with people sharing music through peer-to-peer networks. But music is just one example of documents that you can share through these types of systems, and a large part of why this is happening is people inadvertently share other types of documents that they're not supposed to.

Julia Allen: So what is it about the peer-to-peer networking environment that allows things to be shared that you might not have intended to share in the first place?

Scott Dynes: But I think the reason that it works here is because people are unfamiliar with the technology. I think there are examples where you can share either a directory where you actually want to share particular stuff, or you can share your entire hard drive which inadvertently will be

exposing things that you don't really want to share. But I think the reason is because of the unfamiliarity that people have with the technology, particularly those types of people who will actually be in businesses. Kids, I think, are probably much more adept at this and understand what's going on.

Julia Allen: So is there something in the installation of the client software that ends up revealing or disclosing or allowing access to hard drive contents? Is that what happens?

Eric Johnson: Well there's a number of different factors. Many of the clients come with wizards that allow you to expose specific directories, specific shared directories. But what is often unclear to users is that if you share a directory, a parent directory, every sub-directory underneath that often also is shared.

Likewise many of the clients are designed to really encourage you to share. You can imagine a peer-to-peer file sharing network is not very effective if you have lots and lots of free riders, people who are just looking to download music but not willing to share. And so many of the client developers have created incentives. They'll offer faster service, better search results, larger networks to operate in, if you are sharing a substantial number of files.

Julia Allen: So Eric, to make this a little more tangible for our listening audience, could you describe a few recent examples of cases that you've seen and some of the types of data that show up in this network space?

Eric Johnson: One of the most common sources is home use computing. And we see lots and lots of work documents following people home and then being disclosed. Probably one of the ones I've found most unusual and maybe funny, in a sad way, is during the [U.S.] House [of Representatives] hearing this summer on the subject, I was sitting next to the CIO of the Department of Transportation, U.S. Department of Transportation, and he had to explain to Congress why his chief privacy officer had leaked documents, in this case documents that disclosed employee personnel information. She had brought documents home, had them on a home computer, and a teenager at home had been using a file sharing network.

Julia Allen: My goodness. I think we were corresponding about a case where Citigroup, their mortgage company, again as you just described, someone took home or had access at home to names and Social Security numbers and mortgage information for thousands of their customers, and inadvertently showed up on the network. And I believe you were involved in that case.

Eric Johnson: Yes, and I think that case, as many of these, show the problems that start arising when home and work, personal life and work life, merge. And of course that does for all of us, in varying degrees, and Web 2.0 in many ways is accelerating that. And so you have people who have much of their life on a laptop, both home and work, and when those collide often that's when we see these types of disclosures.

Julia Allen: So Scott why are these types of file sharing networks so hard to control? Can't we use our perimeter protections like firewalls to block outgoing and incoming messages across this network?

Scott Dynes: Well I think they are. And Eric, correct me if I'm wrong, but I do think that companies do try to control the ports that file sharing or peer-to-peer clients use. But I think the peer-to-peer clients have taken some pains to try to make sure that even if there are firewalls that they can - that you can share files with others. And I think, as Eric was saying, a lot of this happens is

because work travels home, and you don't have the same types of firewalls set up at home as you might at your business.

Eric Johnson: Absolutely, or in a Starbucks or in a hotel, anywhere along the business path.

Julia Allen: So what I hear you saying is even though you might have control in your organizational environment or in the network that the organization controls, when you've got this migration, as you said, between home or work or when people are on business travel, the sharing or the control protections aren't nearly as robust.

Part 2: How Can I Find Out if a Peer-to-Peer Disclosure Has Occurred?

Julia Allen: So Eric why haven't we heard more about this? It's still a fairly unknown area and I think some organizations are getting surprised.

Eric Johnson: I think that the very nature of inadvertent disclosures often makes them less visible, and particularly home users often don't even realize that they have been or are disclosing information until it surfaces somewhere, and in the case of some business that might make a media story. But many times what we'll observe is home users disclosing information, for example, about themselves, financial information that could easily and often does lead to identity theft.

But I think in many, many cases the users probably would not attribute it to a loss over a peer-to-peer file sharing network. They would just as equally likely believe that they had lost their credit card number at the Chinese restaurant they ate at two weeks ago or at the hotel they stayed at the week before. And so it's often not obvious to them that disclosure occurred in that way.

Julia Allen: So Scott have you seen or started to see any cases where the loss of sensitive customer information in the peer-to-peer network is actually showing up as a compliance violation or where there might be a legal proceeding launched against a company for insufficient control? Have you seen any of that yet?

Scott Dynes: I've not seen that. Eric, have you heard anything about that?

Eric Johnson: The issues around SOX [Sarbanes Oxley] compliance now is beginning to spend — the auditors are spending more time on the role of information access within organizations, and whether or not certain privileged information is easily floated around within the organization. And so that certainly has and is becoming a bigger issue. But a specific case where a SOX violation or some other violation occurred because a computer was identified as on the peer-to-peer, or any particular website for that matter, I don't know of any.

Julia Allen: Yes, because it seems to me with the proliferation of state-based breach notification laws, and obviously the consideration of a federal law, that peer-to-peer just ends up being another mechanism or method by which data may be breached, and therefore you start into the whole notification process.

Eric Johnson: Which of course has happened and this year, earlier this summer we saw Pfizer disclosing that they had had a peer-to-peer loss, again affecting the personal information of a number of their employees, and they had to go through the notification process there. But the fact is there are many different ways that breaches occur and peer-to-peer is just one, maybe one that's not very well understood at this moment.

Julia Allen: So given all of this how can a business leader determine if their organization's sensitive information is out on the peer-to-peer network? How can they find out and try to get their hands around that?

Eric Johnson: Well there are a couple of different things business leaders can do. First of all there are services, there are commercially available services that will in fact monitor peer-to-peer, very much like services that monitor the Internet for brand or other copyright violations.

Of course the issue with that is that at any point in time you're kind of getting a snapshot. And unlike the web where maybe if you're looking for someone that's violating your brand, they're likely to be perpetually violating your brand; that is they've built a website that encroaches on your brand in some way and that website is available 24/7.

In the peer-to-peer it's much different. At any one moment there are millions, in fact 10 million one estimate shows, simultaneous users of the peer-to-peer, and they're all sharing. But a moment later the mix of users has changed — some have logged on, some have logged off — and so documents or other files that are available at one moment will change from moment to moment and from day to day, month to month, year to year. And that makes, of course, the surveillance problem more challenging and one that requires a more ongoing effort.

The other approach, of course, is not just looking to see what got out of the barn but also putting good measures in place to prevent the horses from escaping the barn. And there are many approaches there that firms are actively using, many firms, including good content management practices and software, building strong access control, and also blocking and preventing corporate users from using these types of services.

Julia Allen: Yes, and I would think that clearly awareness and training, education of your employees and network users goes hand in hand with that, so that they understand what some of the risks are, particularly if they're telecommuting or traveling or working from home.

Eric Johnson: Yes, absolutely, and I really should have mentioned that first. We argue that education is a real key to this problem, as it has been to many other problems, particularly for consumers. I think that many consumers now are fairly well educated on the dangers of phishing and have learned to avoid that in many cases. But something like peer-to-peer is still something that most users don't understand.

Part 3: The Upside and the Downside of Peer-to-Peer

Julia Allen: So you contract with a monitoring service and you find out that some of your key corporate documents or information is starting to show up in the network, is there any recovery action or are you just kind of stuck?

Scott Dynes: I think recovery is actually quite difficult. What happens if somebody does share a document, that document will propagate throughout the peer-to-peer network. So not only will it be present on the initial leaker's device but over time it will show up on several devices, over which you have very little control other than kind of sending an email saying "please take this off."

I think we did an experiment — Eric didn't we? — where we put out some documents and saw them propagate over the network, if I recall?

Eric Johnson: Absolutely. And those documents can move quite quickly and be leaked over and over again.

Julia Allen: In fact Eric, I think I read in one of your papers that you actually put up a debit card with a small balance and then measured the amount of time before that balance went to zero — do I have that right?

Eric Johnson: Absolutely. And those kind of honey-pot experiments, both with phone cards and Visa cards and so forth clearly showed that not only are people taking things but they're quite willing to use them. And the threat is a global one, as we saw in both cases, some of the takers being outside the U.S.

Scott Dynes: Yes that's true. And the interesting thing in this case is that it was on the LimeWire network and the fellow who took the card used it to buy LimeWire Pro. I guess he or she really liked the fact that she could go out and find money.

Julia Allen: Well there is some irony there — right?

Scott Dynes: Yeah you can ask the question of Eric, it sounds like P2P has, is getting a bad rap. Are there any good uses to P2P?

Eric Johnson: Well, it is kind of funny that you've mentioned that because I've had a number of emails over the last month or so of various software firms that are using P2P for legitimate positive content distribution strategies and so forth including Microsoft and others that are creating large scale industrial versions of P2P to allow rapid dissemination of information even within a firm or outside of a firm.

Scott Dynes: Right.

Eric Johnson: And in the press right now, they all get kind of lumped together, and I think for those people they feel like that the music sharing has given them a bit of a bad name.

Julia Allen: Like any other technology there's always the upside that makes it attractive and easy to use and makes their life easier, and then with all that, which we typically don't anticipate in advance, are the risks and the downsides.

Are there, in addition to your own papers and publications and obviously the work that the Tuck School of Business and your Center is doing, are there some other places where our listeners can learn more about this subject?

Eric Johnson: Well the U.S. Patent Office has written a very comprehensive report that examines the user interface issues of many of the clients and kind of some of the deceptive or less than transparent features of these clients, and I would recommend that as a great source of supplementary reading.

Julia Allen: I do want to ask you about that though. You raise an interesting point Eric. Is there a way that a user, either a home user or a business user, can find out if one of these clients is installed on their machine? Because I've also — I have to be careful, I know enough here to be dangerous, that actually some of these clients can be installed on your machine without your even knowing it?

Eric Johnson: Well I think in most cases it's pretty hard for the client to get there without you ever knowing that it was there. That said, once it's there it is often not clear that you are sharing while you are sharing. And there are users who have believed that they felt like they minimized a window

or shut a client down that they were no longer sharing, when in fact they still were sharing. Again I think I would reference this report from the Patent Office because it actually looks at some very specific clients and discusses some of the features of those pieces of software.

But that said, again the real challenge here is it's very much a moving target. There are many different clients. There are some very popular ones that lots of people have heard of, but there are many, many clients that are available out on the web, and they are changing. So it really is a user beware.

And one of the pieces of advice I give, just even home users, is that machines that you use for sensitive information, financial information and so forth, it's better to keep those machines from your areas of routine web surfing and downloading, kind of the green/red machine concept, the one family machine that everyone drives the Internet on. Don't keep your bank records on that machine.

Julia Allen: Well again thank you both so much for your time. And I'm sure our listeners will benefit from the education that you've provided, and hopefully we'll get to talk again on another subject soon.

Scott Dynes: Well thank you very much.

Eric Johnson: Thank you.