**Indicators and Controls for Mitigating Insider Threat**
**Transcript**

## Part 1: Deriving Candidate Indicators from Over 500 Cases

**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 operational resilience and software assurance. Today I'm pleased to welcome my colleague, Mike Hanley.

Mike is currently leading the laboratory efforts for CERT's Insider Threat Team. We'll be talking about those today. And Mike and I will also be discussing his team's current research in how to design effective controls and indicators to prevent, detect, and respond to insider attacks indicators and controls that organizations are finding helpful to get a handle on how to manage their insider threat.

Just as a point of information, we have posted three previous podcasts on CERT's insider threat work that our listeners might want to check out. So with no further ado, welcome Mike. Glad to have you with us today.

**Mike Hanley:** Thanks, Julia. Glad to be here.

**Julia Allen:** Okay, so it's been a while since we've had an update on the insider threat work to our podcast series. I think the last one was August of 2009. So could you please give our listeners an update on some of the highlights of your team's work since we last talked?

**Mike Hanley:** Sure. So last time you were talking with Dawn Cappelli, who's the manager of the Insider Threat Team here at CERT. And one of the first numbers that people always ask about is how many cases are we operating from? So all of our research is going to be borne primarily out of the case library that we've been collecting since our original study with the Secret Service starting back on 2001.

And when you talked with Dawn last time, we were operating with about 280 cases of insider threat and now we're up over 500. So that's 500 actual cases of insider threat, crossing IT sabotage, fraud, theft of intellectual property, and espionage involving national security information.

So we've gone quite a bit up in terms of how many cases we're actually working from and that's really helped us out in terms of doing a lot of our analysis and modeling work.

**Julia Allen:** Mike, are those cases mostly, again, from your relationship with Secret Service or do they come from a wide variety of sources?

**Mike Hanley:** So roughly the first 150 came from the original Secret Service study. Since then we've done a lot of collection on our own, just from open source sources. So we do track things like media reports and court documents for current cases of insider threat. We try to track those to a conclusion. And once they've been prosecuted, we'll go through and finish up our analysis and make sure those get catalogued, along with the rest of our insider threat cases.

But the number of cases that we originally got from the Secret Service is now close to about one third of the cases, whereas it used to be the original all starting points. We're getting a lot of stuff from open source.

**Julia Allen:** Great, and what about some other areas of work that have been going on in the team?

**Mike Hanley:** Sure, so we've also continued a lot of our system dynamics modeling work. On our website, we've got the original IT sabotage model which we did several years ago. But we've also recently put out two system dynamics models for theft of intellectual property. So that's a hot topic right now.

People are concerned about, “How can I protect my IP from malicious insider who might steal it for financial gain?” And we've actually put out two models in a single paper, one of which is an “entitled independence.” So this is an insider who feels like they have ownership over an IP by virtue of how much time they've spent in developing that product. And then we've also got an “ambitious leader” model. So that's somebody who's a strong individual recruiting other insiders or outsiders to facilitate their theft of IP that they may not directly have access to, or may want but may not fully understand by virtue of not having been intimately involved in the development of that IP. So that's our more current modeling research.

And we're also going back and taking a look at a study we did back in 2006, where we compared our IT sabotage model to a few cases of national security espionage. So, as I mentioned our case numbers have gone way up from our 280 last time we talked. And we're now taking another look at national security espionage cases -- of which we have over 120 and looking at a new system dynamics model in that space as well.

**Julia Allen:** Excellent, excellent, thank you. So as I was getting ready to talk with you today, I reviewed a recent paper that you've developed on deriving technical controls and indicators of insider attack. And I noticed in the paper that you said that when you're analyzing a crime committed by an insider, you often have to deal with incomplete data.

You've got pieces of the puzzle. So when you're putting a case together, to make it a legitimate case for your case database, what are some of the essential information pieces or items that you look for?

**Mike Hanley:** Sure, so particularly from my work, so in the technical space -- looking at ways we can apply our knowledge of insider actions against organizations to technical tool sets and improving ability to detect these events we're looking at things like what types of assets were targeted by the insider? So do we have some idea what types of things are being targeted and why? Are they always assets of value? Are they business plans? Are they engineering specs? Does it vary by type of job that the insider is in? -- things of that nature.

We always look for the type of exfiltration involved. If it's a theft of IP case. So is it something where the insider printed a document and walked out the front door with it? Or was it sent off in an encrypted zip file off the network? So we try to distinguish between those because obviously they're going to have very different indicators that we might look for in terms of the detection.

We also look for who in the case of an asset being delivered to someone else -- we look for who's the intended recipient. Does that tell us something that's of interest about the case? Is it a competing firm? Is it something I'm just taking for myself to try to sell independently? Those are some things that we also try to track. So we try to track a range of technical items that are of interest to us. But those are some of the key ones that we want to make sure we always get access to.

**Julia Allen:** Okay, so do you have an example to illustrate the points you made?

**Mike Hanley:** Yes, sure. So a recent case that I can give you an example of that we've looked at involves an insider who is in an engineering role for a firm that was manufacturing consumer electronics. They had been contacted by the chairman of a foreign firm about a possibly job opportunity. And this was over company email. So they were exchanging notes about, "Yeah, I'm interested in this job opportunity. What would I have to do to get this job with the foreign firm?" And the answer, essentially was, “Bring with us IP that you have access to now at your current job.”

The insider downloaded a very large volume of data within a 30 day window prior to their resignation (which I'll talk about in a second), downloaded a large amount of IP, and sent it off the network through company email off to this outside competitor organization.

So there are a lot of indicators that we try to extract from a case like that. So first off, we know there was a large volume of data moved across the network internally prior to the actual exfiltration of the data. So we can use various network sensors and instrumentation to observe that going on within a network enclave. And see, “Okay, this client workstation is pulling down a very large volume of data or an abnormal amount of data,” if we do some testing to see what the normal base line is, “from a sensitive file server that contains sensitive documents or specifications for one of our products.”

So that may be a first candidate indicator that we might want to look at Another one is, “Who am I communicating with?” So if I'm communicating with a firm that's a direct competitor when it's outside of the country and I have to worry about things like ITAR (International Traffic in Arms Regulations) restrictions then that's another thing that I can query on. And I can look back and say, “Well, I know a lot of people are stealing information via email. Can I tie all these attributes together and look for e- mail that's going to foreign competitors or outside the country with large attachments on it?”

So we can start to build progressively from the descriptors that we see in the cases to start putting together these candidate indicator sets that we deploy on various sensor systems to see how effective they are. So there's an important distinction here. So I use the word “candidate” because these are all very first cut at doing technical work in insider threat. So this is new to us in the last couple of months because traditionally, our work has been behavioral. And we've looked at insider threat as a holistic problem.

Now we're trying to focus in a little bit on applying some of that knowledge to these specific technical problems. And to do that well, that requires working at the intersection of the behavioral and the technical. So taking tools that we've got today or tools that are specifically designed to combat insider threat and focusing on what those are trying to collect or alert on based on things that we learned from our behavioral models.

So I mentioned in that case example, that the insider was stealing the data within 30 days of their resignation. So that's one of the more striking findings from our theft of IP model that I mentioned earlier, is that something like 65 percent of our insiders steal that data within a 30 day window. So that's a key finding that I can alert on. And based on what we know about the insiders stealing IP, that's something that gives us a basis to run with.

So when we start wrapping these elements up together, we end up with a complete candidate control set that we can start testing and tuning to see if this is something that might be effective general guidance to distribute out to people.

**Part 2: Prevent, Detect, and Respond; CERT's Insider Threat Lab**

**Julia Allen:** So Mike, you've started to get us into some of the content around indicators and controls and profiling behavior in some of these cases. But I think it would be helpful, so it doesn't feel too overwhelming, to think about them in categories as you describe in your paper.

So you talk about prevention controls, detection controls, and controls to help with response. So could you say a little bit more about both controls and indicators in those three categories or buckets?

**Mike Hanley:** Sure, so any candidate control or indicator that we look at generally is going to fall into one of those three buckets. So the first that we look at are what are candidate controls or indicator sets that might assist with preventing the insider from completing their crime? So these would be things that would be more of an active defense that would prevent an insider from, say, e- mailing something off of a network or downloading a large volume of data from a sensitive file server. So this is going to actually have the intent of preventing the insider from completing their malicious act.

One of the things I think we're doing a good job looking at right now is detection. So we know enough from our system dynamics models about how insiders behave. And we know enough about the technical details of how they go about stealing information or damaging IT assets to say, based on matching between those patterns of behavioral and the patterns of technical actions, that if you see these signs this is an indication of malicious behavior that you should potentially look at.

So detection being we think something happened or we think something is in progress that's worth looking at to determine whether or not it is actually a real instance of a malicious insider damaging your organization or stealing data.

**Julia Allen:** So when you're in the detection mode, can you give an example of some of the --, I don't really want to so much get into tools but how do you monitor for supporting detection? What are some of the key methods that can help an organization know if an insider attack is in progress or is about to happen?

**Mike Hanley:** Sure, so I think first things first to step back from the technical side and recognize that insider threat is not just a technical problem. So you can intervene early or have some indication that somebody might be at risk of committing a malicious act against your organization if you have good practices across physical security, human resources, and individual managers to get buy in across the organization. Insider threats are a real problem and if we're all talking about things that we think might be suspicious, then we can intervene before these things ever become a problem.

From a technical side, though, a lot of organizations have a robust set of network instrumentation already in place. And I think one of the things we find is that people don't necessarily realize that they can use a lot of these tools that they already have so things like packet capture, IDS (Intrusion Detection Systems), any other kind of firewall architecture or any kind of host -based system that potentially could observe a lot of the indicators that we see in our insider threat database.

For example, if I see somebody sending off a large volume of information off the network and I'm concerned that that could be a problem, well I could maybe use my mail server logs to go back and see who's sending off large attachments and who are they sending them to. And I might find in that data, “Oh, I have an employee who's sending off large attachments that are encrypted to a foreign firm.” And that might be something I want to go back and investigate.

Another example might be if I'm downloading a large volume of data from an internal server that I shouldn't be downloading from, I can potentially see that with normal network instrumentation that suddenly shows there's a huge amount of utilization on this link between a client and a sensitive server. That's something I should go maybe look back at.

And the more of those indicators that you can tie together, the more confident that you can be that these candidate indicators are maybe leading you toward somebody who is actually behaving in a malicious way.

**Julia Allen:** Okay, so we've talked about prevention and detection, so all well and good but bad things still happen. So what about controls and indicators on the response side?

**Mike Hanley:** Of course insider threats are a hard problem. You may not always be able to prevent and you may not always see an attack that's happened to your organization. So if you find out after the fact, a lot of the candidate controls and indicators that we want to deploy have the added benefit of assisting incident responders.

So they're going to leave behind trace indicators or trace data that will help us go back and reconstruct the types reconstruct the events of the crime that the insider perpetrated, to help us figure out, “Okay, who was it that did this? What did they take? And what's the damage for our organization?” Because we find that, in a lot of feedback that we get, a lot of organizations don't prosecute insiders not just because they don't want to be embarrassed but because they don't have enough information to give back and confidently say, “It was so-and-so that actually committed this crime.” So we want to help people get back to having some sort of attribution, so that if they do choose to prosecute, they've got that option.

And I think the more information that you've got with respect to how a crime occurred, hopefully that can help you facilitate your recovery operation and get back to a steady state.

**Julia Allen:** Excellent, excellent. Well let's -- I'll tell you what, let's turn our attention to, I know, an area that you're quite passionate about, which is the insider threat lab. And so you've got this incredible wealth of cases. You're starting to get a real handle on controls and indicators and then controls at the prevention, detection and response levels. But I know that the lab is all about putting this all together in terms of real cases. And I know you've developed scenarios that you actually use in your courses to help people walk through and experience real cases.

So can you say a little bit about how you actually put a demonstration together in a lab setting, to explore the extent to which indicators and controls will help in a particular attack situation?

**Mike Hanley:** Sure, so the first thing is there's a direct benefit to people who are seeing some of the demonstrations that we're putting together which is that I think it does a nice job tying back behavioral modeling, which we're doing, which can sometimes be seen as a harder thing to understand, that's maybe a little bit more of an academic exercise ties that back to tools and operational problems that people are having in organizations by showing to them, “These are the patterns that we would observe in a real case, that we can recreate in our lab. And these are how the indicators, the technical indicators that we're concerned about might be instantiated in such a way that you could observe them.”

So it provides people with a really good learning aid that takes them through the steps in a very methodical way and shows them what they should actually be looking for. And maybe sets off a light bulb that says, “You know, maybe I could maybe be doing something like this with tool X, Y, Z that I already have deployed in my organization. Or I've seen that type of pattern before. This is the kind of thing that I maybe want to keep looking at in my organization.” So it does a good job driving a lot of the messages home when you show somebody a demonstration of a case and what could have been done to prevent it.

So we look at things like an insider who's emailing sensitive IP off the network to a foreign organization. So we look at what are some things that we would want to include in a demo that demonstrates that well. So things like technical indicators of sending email to a competitor domain. That's something that we can alert on and that we can actually show in a demonstration.

But from a behavioral side, insiders who steal IP within a 30 day window, that's again, that's one of our key findings from our theft of IP model that 65 percent of insiders do that. If we can show that in the demonstration there are real timestamps that show seven days before the insider left, they were sending off email to the foreign competitor containing very large attachments.

We also show how the organizations can use the tools that they've got today. So one of the demonstrations we recently did, we used Splunk, which is a centralized log aggregator that we've seen in a lot of infrastructures.

So we can say, “Here's how you could do maybe a simple example query that doesn't necessarily fit your organization exactly. But we want to give you this example and show you how you might be able to tailor it to the tools that you've got or the logs that you're feeding in from your own systems. And write a query that might help you roll up all of these indicators into one set.”

So it ties, like I said, it ties together the behavioral indicators, the tool sets, and shows you how you can roll that all up into a very good operational way if you're working in, for example, a security operation center or in a NOC and you're worried about insider threat.

## Part 3: Key Roles for Using Indicators; Future Research Directions

**Julia Allen:** Great, well thank you for that explanation. It makes me want to take one of your courses and sit through one of the demonstrations. I think it would be really a very educational experience to immerse yourself in one of these situations.

I did want to ask you, though, what roles in the organization are typically involved? I mean, clearly, your IT Ops folks are involved, maybe some of your information security folks are involved. But when you talk about these indicators and controls, who's actually using them or benefiting from them or trying to monitor for them? Who do you find most typically involved?

**Mike Hanley:** So I think the focus that we're trying to get with these controls is we're looking at folks who are working at security operation centers. So these are the technical staff who are monitoring all of the same sensors that you normally are worried about today that you think of in like a typical border architecture. So I'm looking at what's my firewall telling me? What's my log aggregator telling me? What are my intrusion detection systems telling me?

By and large, we're looking mostly toward the more technical operators who are running these systems, who are looking for these types of alerts on a daily basis. And teaching them how to use the same tools that they're maybe using today, in a little bit different perspective to look not just at what's coming in from the outside of the network but what's potentially going out or what's happening solely within the enclave that's concerning.

**Julia Allen:** Well this has been great, Mike, in terms of just setting the stage and giving us a refresh and an update on just some, I think some very foundational work. I know that there's a great deal of interest in because so much of what you're doing is based on actual cases, too, so that people can learn from the research your team has done.

So, speaking of which, what's next in your pipeline? What do you have planned for the coming year in terms of your research efforts? And also if you could point our listeners to some sources where they could learn more about this work, that would be great.

**Mike Hanley:** Sure, so the we're going to take this lab and we're really trying to address the problem that I see in the insider threat space, which is that we have good tools, both specifically designed for insider threat detection and response, and also good network security tools that aren't necessarily designed for that but could be repurposed to detect some of these indicators. And we also have really good behavioral models of insider threat coming out of CERT and other organizations. And we need to work at the intersection of those two spaces.

So we need to find ways to map back the capabilities of some of the technical tool sets to the patterns that we're observing from a behavioral perspective. So that's where we're really going to be driving a lot of our work over the course of the next couple of years. Like I said, one of the ways we're doing that is we're taking our really large case database sets.

So we've got over 500 cases. We break those down even further and look at some of the things that I mentioned earlier. So what types of assets are being taken? How can we map that back to a given behavior? And can we build a really good set of standard controls, maybe, that we've tested thoroughly with some pilot organizations that we’re looking at testing with over the course of the next couple of months to vet the controls, provide good guidance on how to customize them, based on what types of tool suites? Because, again, we don't want to be focused specifically on a given tool. We want to abstract ourselves from that a little bit and then provide guidance on how an organization can customize based on specific intelligence, operational needs, other concerns that they want to address.

So providing that standard set with good guidance on how they can implement these controls. And not just help organizations by saying, “Here's what we know.” But help them become a little bit more self sufficient by saying, “These are the things that we recommend. And these are things that make sense to us based on our data.”

But clearly we don't have all cases of insider threat. A lot of the cases go unreported, a lot of cases aren’t detected, so the cases that we have are the guys who got caught. And that's why we say these are candidate controls and indicators right now because they're really based on that large but limited subset.

So we want to make sure that people have the ability to take their own cases, apply that knowledge based on our behavioral models and our technical guidance, and then develop their own candidate controls.

We've got a paper out that describes some of these things. As you mentioned earlier, you alluded to a paper that we just wrote. It's called, *Deriving Candidate* -- *Deriving Candidate Technical Controls and Indicators of Insider Attacks from Socio-Technical Models and Data*. That's a mouthful but it's essentially how do you map back behavioral data to technical controls. That's going to be available on our Insider Threat website on CERT.

**Julia Allen:** Excellent, excellent. Well, Mike, I can't thank you enough for your time, and great advice, and the work that you and your team are doing.

I think our listeners will find this invaluable, and so I thank you very much today.

**Mike Hanley:** Thanks, Julia, for having me.