How to Be a Network Traffic Analyst

featuring Tim Shimeall and Timur Snoke as Interviewed by Suzanne Miller

---

Suzanne Miller: Welcome to the SEI Podcast Series, a production of Carnegie Mellon University's Software Engineering Institute (SEI). The SEI is a federally funded research and development center sponsored by the U.S. Department of Defense. Today's podcast will be available on the SEI website at sei.cmu.edu/podcasts.

My name is Suzanne Miller. I am a principal researcher here at the SEI. Today, I am pleased to introduce Timur Snoke and Tim Shimeall who are part of our CERT Division. We are going to talk today about how to be a network traffic analyst. So those of you out there looking for a career change can listen and see if you like what you hear, and we can talk about what it means to be that. But before we get into that, Timur first and then Tim, tell us a little bit about: How did you get here? What is it that brought you to the SEI and to this work in network traffic analysis?

Timur Snoke: I came to the SEI after a long road working everywhere from washing pots and pans on a dude ranch to being a high school teacher. I learned that I enjoyed understanding how things work and understanding at a deeper level of how things are connected and how things work together. It also becomes very interesting to find out when things aren't working, why they aren't working. So I went from being in a high school classroom to working for the school district to doing technology things to eventually working as a consultant or a contractor for businesses and organizations in a wide variety of industries. I had the opportunity to come here to work at the SEI, and I jumped on it.

Suzanne: Cool. And you had no idea what you were getting yourself into, but now you're a network traffic analyst. And Tim…

Tim Shimeall: Well, I was a tenured faculty member working for the federal government out in Monterrey, California. I managed to come out here for a sabbatical for about five months, doing some studies with the CERT Division here at the SEI, and found the environment attractive enough that I resigned from tenure and came out here, principally because I'm a very data-oriented person and I very much like to see the interplay of data that's in place.

So I swapped Monterrey for Pittsburgh and have occasionally had my sanity questioned, particularly in February. But it's been 20 years now, and I've really found the SEI to be a tremendously wonderful place to work. Along the way, I have been able to do some really, really, really interesting things including looking at a large amount of network traffic in a wide variety of environments including the 2002 Olympic Games.

Suzanne: Oh, you got to be part of that. Yes, that was a very big deal when that was going on. So like you, I left California in 1993 and came here. Yes, people do question my sanity, but I do the trick of going home to visit my family in December, so that sort of helps a little bit. But, all right, so very different backgrounds, but both ended up in sort of the same arena. First of all, tell me: What is a network traffic analyst and what do they do?

Tim: Well, network traffic analysts, as the name implies, look at network traffic. It's not necessarily just one form of network traffic. There are several distinct types of data that analysts may look at including things like, routing data—*how is the information being sent from one location to another, what paths are currently involved?*; packet data—that is, *what units of information are being sent back and forth between computers across various networks?*. Looking at packet data could be traces, which is basically packets where you strip out the content and just keep the header but still retain packet-by-packet information, or a flow where you assemble and summarize whole collections of packets that correspond to a connection between one computer and another.

Suzanne: Why do we care about all of that?

Tim: Largely because it can give us some understanding of what's really happening on our network, an unbiased understanding, so I'm not dealing strictly with things that I initially believe are wrong, but rather, I can look and see what's on the wire, what's moving, how much bandwidth is being consumed by which applications. Do I see the types of traffic I expect to see that indicates services are still alive, or is anything there that might indicate an attack?

Suzanne: So I think coming from CERT, we make the assumption that a lot of what we are looking at when we do any of these analyses are things that might be leading us to understand vulnerabilities better, leading us to understand the attack: *What's the attack surface? What is the attack profile that is available within this network environment?* all those kinds of things. Do we also use network traffic analysis for other kinds of technology understanding? I'm thinking about things like: How close are we to needing to add resources within the cloud environment that we're supporting, or things like that.

Timur: In addition to what you were talking about—capacity planning with transitioning to cloud or something like that—we also are looking at the utilization of the network connections in

between different devices to figure out *Do we have enough capacity to let our applications run with optimal performance?* We are also monitoring how the applications are talking to each other and what applications we have present on our network. Frequently we find that the environment that we have that creates that attack surface is unknown. When we go into a customer's network that is in the midst of an incident, and we ask them the question, *How many web servers do you have?* There are few that can actually answer that accurately. So often this network traffic analysis is just to provide situational awareness to understand: *What is the baseline of the environment that we're trying to defend?*

Suzanne: What should the baseline be and what is it? Then, what is the variance? So that we can understand that there might be something good or bad happening.

Timur: Right, often we find that the network traffic analysts wear two hats where there are network security and are also more of an IT service-support capacity.

Suzanne: OK. What is it that makes a good analyst? I will tell you I don't think you are going to bring out qualities that are part of my profile. I like to know how things work, but it is more that I like to know how the organization works and things like that, more than understanding all the bits and bytes of data. What is it that makes a good quality analyst that can deal with this kind of data on an ongoing basis?

Timur: Part of it is the ability to use a wide variety of tools to answer questions about what is happening on the network and to figure out ways to go past inference and supposition and to get facts that can actually provide support for the hypothesis that you are coming up with.

Tim: Some of it is what you have been talking about. The analyst needs to be able to keep a higher perspective and to apply insight. *So it is not just these bits and bytes are moving across the wire, but it's this kind of interaction that is important to my organization. What's going on with the domain name server? How much of the resolutions that are being made are business-relevant versus not?* Sometimes the non-business relevants are things we're going to tolerate, recreational use within allowable parameters and things like that. Sometimes, if you are starting to see a huge uptick in resolutions going out of your organization, that can be a sign that there's malware there, because these manufacture domains that are deliberately in design not to be present already in the DNS cache. They also have to have the ability to ask meaningful questions of the data. There's so much data that's out there.

Suzanne: That's what intimidates me, the overload.

Tim: There's huge volumes that are there, and so you have to come into it with some understanding of: *This is what I'm looking for, and this is why I'm looking for that. What does it mean to my organization? What does it mean to the computer systems? What does it mean with*

respect to security? And [a network traffic analyst has to] be able to have that insight when you're looking at the data because that helps to scope and bound your data as you move forward.

Timur: To add to that, sadly, there is really little formalism with what a network traffic analyst would do. What they have to do is to bring their own skills to bear, and that might be drawing on things like a knowledge of statistics or behavioral analysis or maybe some deep learning or machine learning applications.

Suzanne: But fundamentally it is a classic problem-solving skill set, right? I've got to be able to look at an environment, characterize the environment, figure out what is meaningful in that environment, and then match it up to the problem, the incident that I'm trying to deal with.

Timur: Right.

Suzanne: So I've got the divergence of scanning outwards and not pruning the tree prematurely, and then being able to come meaningfully and in a timely way to what's converged as, *Here's the place I really need to look*. People that have those kinds of abilities, adding in all of the specificity of being able to deal with the mountains of data—that's the piece that really just intimidates me is mountains and mountains of data.

Timur: And then to translate that into something that is meaningful to the decision makers; yes, absolutely.

Suzanne: Who are the decision makers that network traffic analysts typically deal with?

Tim: It could be your CISO, chief information security officer. It could be just a network-security reporting infrastructure within your organization. You could be dealing with a security incident and your goal is to get information back to the responder as to which systems were involved, what ways you see those systems involved, things like that.

Timur: Also auditors, compliance people would also have a stake in what a network traffic analyst can provide.

Suzanne: I am even imagining that they might even interact with some of the vendors that are coming to sell the organization, *You have got to have this blah, blah, blah, blah*. That traffic analyst is someone that can say, *Well, we don't really have that kind of thing going on*.

Timur: So to that point, I think one of the things that is incumbent upon somebody in this role is to develop a fluency with the tools that are available to do the work, but more importantly is to understand the underlying technologies that are involved, because many tools provide the same solution. It's just understanding what it is that we are trying to…

Suzanne: The context.

Timur: Right. If I want to be part of this kind of community, and if I'm already part of this kind of community, one of the things I'm hearing you say is: *You have got to keep up*. I know in the tooling environments that I work in, I'm hearing about new tools not quite every week, but almost. What are some of the resources that the SEI supports or that are out there that people who want to be analysts or people who already are analysts can go to get up to speed on what's happening now in the network traffic arena?

Tim: Well, the starting point is to first of all understand what your organization needs. There is some amount of understanding of *what's my infrastructure*? and that kind of thing. And that involves simply starting to gather data, starting to build familiarity with the data and the way the interactions normally work so that you can easily detect *abnormal*, *strange*, possibly *threatening*. *Threatening* is really kind of only a possible, not necessarily a definite at this level. Some of this can be helped by looking at online resources like threat feeds and being able to gather information about what to look for that seems to indicate this kind of attack or that kind of attack. There are several good quality threat feeds that are out there. And then going on and exploring and evaluating on your own, so the more you build your own skill set, the better. That being said, it is nice to sometimes get into an environment where there are several experienced analysts around, ideally a bunch of experienced analysts around, and you can trade expertise. You can talk and find some commonality in what activities are involved. You may be able to see some vendors that have information in place, but not necessarily the flacks of the vendors, but the engineering types of vendors, and really get in, dig, *okay, what does this tool actually do? How does it do it? So I can get some feel for, okay, well, that's similar to what I've already got; I really don't need that, or, wow, that's something new that I want to be able to grab*. FloCon, which is held every January in varying locations, I think the next one's going to be in New Orleans this next January…

Suzanne: Maybe I want to be a network traffic analyst after all.

Tim: It brings together on the close order of 200 participants. The participants tend to be very technical people from a wide variety of different organizations, so we get government types, we get industry types, we get a few academics, we get people from the United States, people from other nations coming to share and to talk about it. There'll be technical presentations, there'll be things like that, and this is the 15th one that we've held.

Suzanne: So you have a community.

Tim: There is an active community there.

Suzanne: That is good for supporting people that are just getting into the field because they can learn who is the guru in this area and that area—Keynotes are a clue for that—and then they also get to meet people that they can connect with after the fact. I know in the Agile community there are several conferences that tend to be the ones that everybody gets together and gets to know who the people are that they can ask different questions from.

Tim: We have one keynoter who is from Netflix who is going to be there in New Orleans. So certainly with their needs and their loads, there's a lot of network analysis that they do. He is going to be talking about some of that. So it should be an active and interesting time.

Timur: I think that the value proposition of something like FloCon is really exceptional because it is not necessarily a vendor-driven event. The history of many of the presentations are still available and are online, so if you want to see what the state of the art is today and understand how it progressed to this, a lot of that has been captured and is available.

Suzanne: Excellent. And that's something we'll have in the transcript. We'll make sure that those links are there.

Tim: It's there on the SEI website.

Suzanne: But also, FloCon has its own website, flocon.org. That's an easy place for people to get an idea of what the depth and breadth are of the kinds of things they cover, and as somebody thinking about becoming a network traffic analyst, this gives you an idea of, *if these are not interesting to you, this is where you might want to get off the airplane*. Or if it is interesting to you, then that's a good way to decide that you want to get a little more involved in these things. What are some things that we're working on right now that are interesting about network traffic? Going from being an analyst to actually being a researcher in this arena, what are some of the interesting things that are happening in your world right now?

Timur: I'm doing a lot of work with open source data, trying to find: What can we infer from the things that are available without a price tag on it? And by watching the traffic that's on the network and watching the way that it relates to the traffic that is outside of our network, we are able to get a lot further faster and close the time that it takes to respond to an event.

Tim: We are also developing formalisms to help us to understand and distinguish between actual information and just inferences that might be influential but they're not really actionable and trying to drive the evolution of data into information and then into action.

Suzanne: So trying to get a handle on this, we have got this big lake of data, and pulling the meaning out of it right now depends pretty much completely on the skill of the analyst and some of the tools that they know they have. You are actually trying to say*: Let's put something around*

that that guides people so that that data lake actually gets a little bit pruned before we even have to apply our own skills to it. Is that correct?

Tim: That's correct, and also to be able to apply some rigor to it so we can make some obvious statements as to what is included in this kind of a method, and not.

Timur: Right, and that's where we're going to start bringing in some of the newer techniques: taking all of the data, trying to sift through the noise using automated tools like the machine-learning algorithms or other kind of behavioral analytics and things like that that we have already touched on. That is a really important part of this, because the volumes of data that our network traffic analysts are having to observe is growing at a remarkable rate, while the population of people doing the work is not. And so we have a lot more data for fewer people, and it's a challenge.

Suzanne: And so automation is one of the things we turn to as we learn more deeply about the space.

Tim: We're also looking at: *How do we better explain how we go about network analysis?* We have a forthcoming handbook which is coming out, the SiLK Analyst's Handbook, and that is dealing very directly on: What are the analytical skills, what are the models of doing the task of analysis, not in a cookie-cutter sort of a way or a recipe sort of way, but more a general mindset.

Suzanne: And competencies.

Tim: And with worked examples so that we can work things through.

Suzanne: So it sounds like this is a field that is on the cusp of becoming much more systematic and definable, so easy for people to decide whether or not they fit or don't fit into this kind of work, and we need people to fit into this kind of work. As you said, the traffic is just increasing, and I don't see any reason why we're going to see a stop to that, so it's just going to get more and more. So if this was of interest to you, then you may want to hit up some of the resources that we'll have in the transcript. Go to the FloCon Conference—it's in New Orleans, after all—and learn more about this area so that you can help us out in understanding what's going on with our networks. Timur, Tim, I want to thank you both for joining us today. This is always fun, to talk about what kinds of things we're doing out in the real world and this definitely fits into that side of things.

So, for our viewers, this podcast is, once again, available on the SEI website at sei.cmu.edu/podcasts. It is also available on the Carnegie Mellon University's iTunes site. It is also available on SEI's very own YouTube channel. Thank you for watching.