uberknowledge

Giovanni Vigna, Lastline

The Adversarial Machine-learning Setting

Giovanni Vigna, Co-founder and CTO of Lastline, talks:

Ashwin Krishnan: [00:00:01] Joining me is Giovanni Vigna, did I get that right?

Giovanni Vigna: [00:00:01] Yeah, absolutely perfect.

Ashwin Krishnan: [00:00:07] So, we're at Black Hat. Giovanni and I have spoken before, but a few things have evolved since last time, and one is about AI.

Giovanni Vigna: [00:00:21] Everybody is doing AI.

Ashwin Krishnan: [00:00:26] Let's talk about what true AI means and why it matters more than ever.

Giovanni Vigna: [00:00:33] So, AI is a very wide term. Most people think of AI as machine learning because it's the most obvious application, but it's a subset of AI. Interestingly, it would be difficult to find a company that hasn't used machine learning well before it became cool. You know, I'm guilty of that. My company Lastline has used machine learning since we started in 2011. Now, we are pushing out the message about what we do with AI because it resonates with

what people want to hear: "What are you doing with AI? What are the advantages of that?"

But I want to take a step back, the concept of machine learning is interesting because many of the algorithms that have been developed to learn have been developed for natural language processing, for image recognition, for voice recognition. Those are all environments in which the data set that you try to model is not fighting back. So, the problem with security is when you try to apply machine learning to malware or malicious web pages or malicious e-mails, the bad guy knows that you're doing it and can do many things, they can morph, steal your models that you created. Research shows that with a few hundred requests from these automated models you can steal what has been learned. Once you've stolen it, you can just morph your stuff to look like that. So this setting, an adversarial machine-learning setting, is very different from the setting in which those basic techniques have been developed. And the problem is that if those techniques are used verbatim in one field to another, they become very easy to evade.

Ashwin Krishnan: [00:02:33] So, I need to ask you a question. I've never heard anybody speak about what the fundamental difference is between applying machine learning to security versus non-security applications. Do you think this is something customers understand? Or is it still at that infancy stage where they are barely getting their heads around data sets and models, supervised versus unsupervised learning, and here you are coming and saying that adversarial machine learning is here, and therefore the models could be stolen and I could interject fake data? …

Giovanni Vigna: [00:03:07] I don't think that people realize that. I think people correctly perceive machine learning as a way to save time. If we think about it, just the two basic concepts of clustering and classification from a data set, for which you have ground truth, are the two ways in which it's used in security and it works really well. You have an event to say these are all the same. You can deal with that with one snap of your fingers instead of having to go through all

of that. While in the other case, you learn what's good and bad from label data and then you don't have to worry about it. Unfortunately, it's not that simple. In an adversarial setting you have many things that can happen. Not only can the people pollute the data set you're learning from, but in addition they can morph, steal, they can do all sorts of attacks. They can even do a desensitization attack. Where they know your model and they send you a lot of benign that will come out as a false positive. And at that point you lose trust in the technology completely. All these things don't happen when you are trying to understand images of cats - the cats are still and are not fighting back. This is one problem.

The other problem, another confusion that I see, and you can see it on the floor at Black Hat, is anomaly detection. A lot of people think that anomaly detection is machine learning and it's not. I would like people to go back to 1986, there was an article by Dorothy Denning about a new technique, instead of writing signature about what's bad, why don't we model what's good, and then everything that is not good is a bad thing. People said, "Oh, that's great." But after the initial enthusiasm in the 1980s, the thing wanes. Why? Because it's really hard to model what's good, because stuff changes all the time, because you cannot cover everything, because it requires a lot of human expertise and human time. And the moment you're finished, you've finally modeled your enterprise, it's already changed.

There is the first problem, the learning of what's normal. Learning is a key word here because anomaly detection was not very popular, but then people realized they could use machine learning to learn those models instead of having somebody manually do them. And that's why anomaly detection supported by machine learning became very popular because, "Hey, I can use those models to learn what's normal and then find the variation. Great, problem solved. Security's done." The problem is that anomaly detection is based on two fundamental assumptions. The first one is that something bad will generate an anomaly. The other one is that something that is anomalous is bad. Both of those can be false. There are attacks that you will not be able to see as an anomaly. Could be one UDP packet where one string is slightly longer, so you have a

buffer overflow in the network application. You will not get that with anomaly detection. And there are other things, you try to do anomaly detection at university, people do everything - you know, one moment writing a paper, then writing some python code, and then running a huge simulation destroying the SIP views of 50 machines. There was no way to actually model that. Even in an enterprise that might have a more stable environment. There is always the guy, there's a deadline, it shows up at 3:00 a.m. in the morning and then you start firing those alerts and suddenly people are discouraged and say this doesn't work, it's too much.

Now we get to the point of what we do at Lastline. We say, "Hey you know that is something that we already use in many forms. We do machine learning, we do all these things, but we now can combine these to save you time." Our different way of doing things is, yes, let's do anomaly detection, but let's use those as events like information, as context, not as detection. Then use our technologies that have very good detection as a grounding mechanism. So I'll be able to tell you, "That machine is compromised," because I have certainty. I have all these events correlated, but then I can use anomalies around that machine to say, "After you got compromised, this happened and this happened and this happened." And I create an anomaly profile and see if it happened in other places of the network. By using this combination, I'm not just throwing anomalies at you. I say, "I'm showing you this anomaly because it's related to something bad and I'm sure it happened. It might still be a false positive as an anomaly, but you want to look at this. Not at everything, you want to look at this."

Ashwin Krishnan: [00:08:18] This is particularly important for CISOs and their teams because the number one commodity that they do not have is time.

Giovanni Vigna: [00:08:31] Correct.

Ashwin Krishnan: [00:08:32] So, while everything is about how do I drive topline or deal with skills shortage, I haven't heard people talk about, "How do I save you time?"

Giovanni Vigna: [00:08:46] Absolutely. And that's what drove the product that we put out. The basic idea is that analysts are a very difficult to obtain commodity, and a difficult to retain commodity; difficult to train and retain. So there is always this idea, "Oh my God, I lost another SOC analyst! I have to find it. They're expensive." You want to optimize their time and their attention. So we did a few studies to see, to say to people, "What is it that you need?" And they replied, "I want to make a decision fast. Most of my time is spent on, I got piece A, piece B. They tell me that they touch the same IP, they're not even related, they are just a coincidence, and then this IP was given to another machine, and I just don't get it." Instead they want something like, "OK, this is something that I give to IT, I preinstall on this machine. This is something that I pass to a higher-level analyst that will do an emergency response to this."

We want to really empower the SOC analysts to make those decisions fast because that is saving time. Saving time means saving money and that's something you can measure. You can measure how many incidents you did last month and how many incidents you handle this month. If you have a good product that should go from 10 to 50, 100, and so forth. It's something that is measurable because security is very difficult to measure. I mean over time, you might be more attacked one month, but over time you will see an improvement if you have the right tools. So, choosing the right tools is really important because it empowers your analysts to do more in less time. Hopefully even do more for a less-skilled analyst, because as you know there is tier one, tier two, tier three. The tier one, they're the first responder and the first pair of eyes on the matter. They have a very important decision to make, they have to right there decide, "OK, we've been breached. This guy is going after our database. We have to immediately shut down this particular thing." Or they look at it and say, "Oh, somebody clicked on a phishing link and downloaded some document with an

exploit and suddenly they are bitcoin mining on his machine. Ok, let's get IT to just reinstall this thing."

Those decisions are really important because you make the wrong decision and you could send an expert to put out a fire that doesn't exist. It's like sending a firefighter to take a little cat off a tree. When you do that too many times the firefighters stop because they have serious things to take care of. But if you make the other mistake and you say, "Oh, this is small change," you'll be hacked. The ability to triage an incident is fundamental to save time.

Ashwin Krishnan: [00:12:13] One of the other takeaways for me from these Black Hat conversations is the fact that there is a sense the adversaries are using institutionalized platforms like phishing as a service. It's almost like everything that the defenders or the software developers have the adversaries are using. So, in some sense are you getting to a point right now where because you have IaaS, PaaS and SaaS available to the adversaries, do you start seeing more homogeneity in anomaly detection itself because there is a large community that is starting to use similar platforms?

Giovanni Vigna: [00:13:07] It is true that the cybercrime community has been commoditized, you have service, specialties, and so forth. I don't think that you will see uniformity, I think that you will see - and you see already on the black market - things that say, "I can bypass these AI engines." It's always a call, you give something to an antivirus and it says yes/no independent of the use of machine learning. And so, they're very smart people that look at this and look at how these tools make decisions, and they say, "OK, how can I morph?" Sometimes you add sections to your binary to reduce entropy so that it is detected in a different way. But there are scientific ways of doing that. So we will see, I think, more of this offering: we beat AI. On the good side, I think that the solution to that is not relying on one single technology. Having one AI approach or saying we don't use signatures we only use AI, that doesn't work.

For example, if I have two classifiers - something that I always find interesting - I have something that can tell if it's malicious or benign. And then I have a machine-learning system that groups together this data. We see a group of documents and we see they're all classified as malicious except one that is classified as benign, but they're very similar. The machine learning told me this cluster is similar. Then I have a decision. That is imperfect because the one that is benign has escaped, has been able to fool the classifier. But now it's classified with all the others, and you can look at this as an invasion. I go back and improve my classifying. If you have these reinforcement loops within machine learning then you can win the war. If you just think, "I'll take this algorithm that I just found on the slides on the web and I'm gonna apply it to malware," that ain't gonna work.

Ashwin Krishnan: [00:15:19] That leads me to my next question. There are two schools of thought. One is the days of many vendors and heterogeneous environment are over because it's too hard for CISOs to keep up. There's going to be mass consolidation and there is going to be 10-20 vendors tops. On the other hand, if you are steeped in trying to mitigate this, as CISO you have to spend the time and the effort. You have to evaluate best of breed while you are fighting time. How does a CISO come to terms with that?

Giovanni Vigna: [00:16:14] I would say that depends on the size of the company because with a small company I just say, "Go to an MSSP," because they even have a security thing. So, they go to an MSSP and say, "Hey, take care of my security." Now, they can work on their tools, but fundamentally the decision is out of their hands.

Ashwin Krishnan: [00:16:36] Have they outsourced moral responsibility at that point in the MSSP example?

Giovanni Vigna: [00:16:42] Yeah, I mean in a certain sense you outsource your ability to see things. What you do with it is still within your realm, your policy, your decision. What we see is the moment that a company becomes bigger, they

don't want to go with a single vendor because there is no single vendor that does everything right. Maybe a vendor has a fantastic endpoint or a fantastic CM but then their analysis on the network is not that good. You want to have an unpredictable combination of technology that makes it harder for somebody to say, "Oh those guys are vendor or X Y Z shop. We know what they have; let's go with it." When we see the sophisticated high-level CISO, they're listening. It's their job to find out the best way to do this and then they know how to create glue. That's why every vendor that is not doing something wrong has APIs so that you can code stuff, take one thing from here put it there, and build your own system. It takes effort, but the benefits that you reap after you do something like that are very high.

Ashwin Krishnan: [00:18:04] So going back to the CISO world, what you are talking about is, you have to have the ability to truly keep up with the technology. But on other hand, you have to take off your tech hat when you're talking to your CMO, your boss, the board. That's pretty hard to do. And if you fix it up and start talking tech to your board, you're just not doing your job. Has the bar for being a CISO suddenly risen?

Giovanni Vigna: [00:18:44] I think a little bit because technology has become a lot more complicated and more difficult to explain. It's funny, you go to a booth of a technology here and read their message, you have no idea what they're doing. It's not like, "We detect malware, or we authenticate people." I think that the important thing there is to take all the tools that you choose to implement your security and define your security policy and report up the results. We have remediated these many incidents. We have blocked this many. So, what you report up are the metrics, the key performance indicators that, as a CISO, you want to say, "Hey this is my job. These are the policies that we have in place. We have this training and we had this thing," and those are things that the board understands. If I go to the board of my company and say, "Hey we have blocked documented, we've blocked these many attacks. There were these many attempts at phishing. We were there and we were able to block them all." It's that goalie situation, "We had penalty kicks and we didn't let one in!"

Ashwin Krishnan: [00:20:08] Yes, go back to the analogy that you just brought up, which is a brilliant one. For the purpose of the listeners, if you're a CISO and your score is 0-0, it's actually the highest point in your career. But to explain how difficult it is to keep a 0-0 score, that's where the smarts of the CISO lies.

Giovanni Vigna: [00:20:29] Absolutely. So these indicators, number of incidents processed, and the pipeline and all that, is a good way to speak upwards. If you tell people, "Hey, we now have better static analysis tools to look at the macro into google docs," people don't understand that. You need to go and explain the impact of what you do in terms of saving time and preventing breaches.

Ashwin Krishnan: [00:21:03] So finally, I know we're probably day four or something like that at Black Hat, what has been an aha moment for you, if any?

Giovanni Vigna: [00:21:12] Good question, an aha moment. There was actually a nice company that I really like and what they do is they look at your screen and you can have certain keywords. They do OCR on your screen. And whenever a certain key word appears they can tag that document and put it in a shared channel. And I like that because that comes from the experience of actual SOC analysts. Everybody wants to use their own tools. If you tell an analyst, "Use this tool." They're like, "No, no, I know how to do my job. I'm using these tools and these tools only," but I want to talk to people using different tools. So, I thought it was a hot idea. Let's forget about all this integration. Let's look at the screen. And I think it's I think it's a fresh idea. I don't see fresh ideas very often.

Ashwin Krishnan: [00:22:13] OK, thank you. Thank you for your time.

Giovanni Vigna: [00:22:15] It was a pleasure.