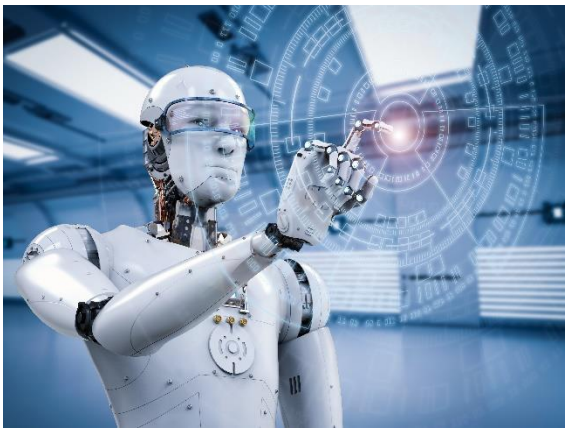
The analysis of security

Justin Richards MBCS spoke to Giovanni Vigna, a professor of computer science at the University of California and co-founder of Last Line, about all aspects of cybersecurity.

What security trends are you currently interested in or worried about?



The security industry is developing techniques based on machine learning and artificial intelligence, but oftentimes these techniques do not take into account what we're trying to learn about, or we're trying to model by using these techniques, and is fighting back.

So, if you think about how machine learning and artificial intelligence were born, they were born to process large amounts of data and recognise, for example, images or the natural voice or text. All these subjects of analysis are not fighting back. A picture is not trying to pretend to be a cat, it's just a picture.

Well instead, when you apply artificial intelligence techniques to programs, to documents that could be malicious, these documents can fight back and can, for example, decide, 'hey, if you're using this particular machine learning technique I can change these few parts and keep from being bad, and will be classified as benign.' So, these are the risks that I think are going to bite us back in two, three, maybe even five years from now.

Are the good guys drawing level with the bad guys or are we losing the battle?

I think that the bad guys will not win the battle, I think the good guys can and will win the battle in my opinion, but we need to be able to use machine learning in an effective way. There is this concept of adversarial machine learning where you have to learn and apply learning modelled in an environment that is fighting back.

It's our responsibility, as good cybersecurity researchers, to develop techniques that are resilient to this kind of attack. So, I think that we're going to lose the battle if we take the techniques that we develop, like image recognition, and text recognition, and we apply those techniques in a naïve way to this new domain.

But, if we are well aware that the domain that we're trying to analyse is fighting back, then we've got a chance to develop novel techniques that are specific for the security domain that can be effective at stopping these criminals.

Do you think we've got enough security professionals?

I think that this is a problem that permeates the whole security field. Creating enough people that can effectively perform security duties of any kind is something that has been a thorn in the side of the government, of organisations, even in education.

We cannot, as a university that has a strong presence in security, produce enough people because there is such a demand. So, this demand is what has driven, in large part, the introduction of artificial intelligence in security because many of the tasks that were carried out by humans they couldn't scale up.

We didn't have enough analysts, enough signature developers, and so forth, to fulfil the demand and so people started looking at automated techniques, machine learning techniques; AI techniques to cover some of these tasks. Of course, this is not a solution.

We will always need a human 'in the look' eventually. But you can try to optimise that time as much as possible. And, in fact, one of the trends in the industry is really a focus on what we call the first level analyst. So, the person who is looking at a screen looking for the first evidence of a possible problem in the network.

IBM's Grady Booch has said that security is partly the responsibility of the software developers, to start thinking in a more security-minded way. What are your thoughts about that?

Oh absolutely. I teach classes in software security development and my point is that we have to start thinking in an oblique way, so instead of thinking 'I have this task, the software has to build this', after you have designed it you have to look at it and say: 'can I abuse it? Can I use this data in a different way? Can I send too much data? Too little data? Data of a different kind? What is my program going to do with this anomalous input?'

So being able to look at this and understand the implications of the data that you didn't foresee is very important because this is exactly what a vulnerability analyst, which will be a hacker that tries to find a vulnerability, tries to do.

When I look at software and I have to find a vulnerability I try to think, 'the developer - what assumptions did he or she make?' And I try to take that assumption and try to break them.

Do you think there should be a bit more of a voluntary code of practice within the IT arena, trying to get people, particularly software developers, thinking more about security, and making it more of a professional code of ethics?

I think so. I think that, for example, at UCSB (University of California, Santa Barbara), we teach an ethical course for engineering, for engineers in the college of engineering, and this is important because every time you build software, in a way, it's like you're building a house or a bridge. You have to understand that your software will be used in certain contexts and you have certain responsibilities about its functionality.

Right now, we don't have laws that determine that the quality must be of a certain kind because it's very difficult to define for an algorithm. We let the market decide. If you develop software that is crappy, people, unless they desperately need it, they won't buy it. But I think it's not the right way to do it. The right way is to teach people that the software has to be developed in the right way and we have responsibilities.

What advantages might a decentralised architecture have?

Centralised systems have been oftentimes looked at as problematic because whenever we have a central system then you could attack that system and take down a whole operation. Because of that decentralised systems were introduced and, if you think about it, the internet was introduced by DARPA (Defense Advanced Research Projects Agency) for exactly that reason. They wanted a network that would sustain, for example, entire cities being annihilated in a nuclear warfare situation. So, having packets being able to be rerouted in a decentralised way was the right way to characterise or to make a network functional in these cases.

So, in decentralised systems you have the advantage that you don't have a single point of failure. However, you have the disadvantage that if somebody controls enough of the nodes it might decide certain properties of the network. For example, they are called byzantine attacks where you control enough of the nodes of a network or the participants in a distributed transaction to create, for example, consensus over things that are not true. So decentralised is good but could also be vulnerable to these attacks. So, pros and cons.

Where are we with mobile security?

I think mobile security has not caught up with the way in which we do security for other systems. For example, the Android eco-system is going a little bit through what the old Windows XP, Windows 7 was going through.

And indeed, there is a core operating system that has maybe good security mechanisms, but then it has to be run on a number of different platforms. Platform here being different types of phones from different manufacturers and each of these phones has a number of chips; little pieces of hardware that handle, for example, the gyroscope or the broadband communication or the audio or the video.

For each of these pieces of hardware, that are different for every phone, you need a driver. The core system might be well-designed and secure, but oftentimes these drivers are developed under a lot of market pressure by developers that are not experts in security.

So, what we have found is that a lot of vulnerabilities are not in the core system, but they are in the driver. The problem is that the driver runs like the core system and so compromising a driver might lead to the complete compromise of the phone in a way that was similar to what was happening with Windows. Now things have changed, there are different ways to load kernel drivers so there is enhancement of the protection, but for phones this is still a problem.

How has the internet of things been causing problems?

The internet of things has introduced a whole new spectrum of vulnerabilities. Mostly because the driver in creating this network of devices is their interaction. Just having Alexa, or whatever other personal assistant, is not very useful if you cannot control the lights, control the speaker, turn on the TV.

Much has been done in order to have these various devices talk to each other. The problem is that these devices are made by different manufacturers with different standards, and so, in order to talk to each other, they try to make them as open as possible.

Openness is a great property because it allows inter-operation, but oftentimes this doesn't go hand-in-hand with security. And so, the resulting problem is that you can have, for example, an Alexa being influenced by a speaker that is under the control of the television that is actually networked and somebody just logged into that and is able to play a sound that tells Alexa to open the door and suddenly the apartment is unlocked.

These interactions are not well understood. We have interactions that we never considered before like vocal interaction, light interaction, temperature interaction. These are not well understood and might cause problems. And oftentimes the firmware, that runs on these devices, is not well-developed and has vulnerability of its own. We have more surface, new interactions, and it will take a while before we master how to have this internet of things in a secure way.

Where are we in terms of traceability of attacks?

Tracing attacks, and not only tracing but attributing attacks, has always been incredibly difficult. Mostly because digital evidence is extremely malleable, it's not something that has molecules attached to it that can leave an actual trace. Sometimes you can spot that certain tactics and techniques have been used - things that are often done by certain groups, in a certain way, and that requires a lot of experience and a lot of observations. After a while you can see that these groups tend to follow certain patterns.

But this is a little bit of a dark art because you have no real way of knowing. In some cases, you might have evidence but it's very difficult to have hard evidence that something has been done in a certain way. So, I would say that you can trace the attacks, you can say where they came from, because you have logging, and you can see where a connection came from, but understanding who's behind it is still very difficult to do.

The National Cyber Security Centre recently warned that the threats of cyber attacks on Britain's critical national security infrastructure from hostile states like Russia, China and North Korea has soared over the last couple of years. Why do you think that there's more nation-on-nation attacks?

I think that nations, especially certain nations, are making major investments in their cyber attacking capabilities. This is because it's something that can provide an enormous amount of benefit with very little engagement. Very little cost as well. You don't have to do a lot of things that would get you caught, that would get you identified and, therefore, it makes sense, from a strategic point of view, to invest in that.

So, it becomes the classic arms race as more nation states develop more security attacks, and the nations that are the targets of these attacks will have to develop better security measures to block these attacks and make them ineffective in some way.

Did you have a mentor that inspired you to get into the industry?

I had several mentors. I would say that I got into security because I love the challenge and I started doing security stuff when I was very young. I just loved security, the concept of, from locks to networks to, 'hey, what's in this computer and why it works that way.'

I did my studies in engineering and then there was one professor that was actually in Santa Barbara that was an expert in security and I really wanted to do a little bit of work with him, so I went to visit him after my PhD, in 1997, and I never made it back. I stayed there, and worked with him for many years, so I would say that he is the reason why I stayed in security for so long.

What are your thoughts about the security levels in blockchain?

Blockchain is fundamentally a way to have a distributed ledger where a group of people can agree that certain things happen in a certain particular order. And in those cases, blockchain is a very useful approach. Blockchain is based on a cryptographic concept that makes it possible to have this blockchain, this ledger, certified.

To me the security problem is born when you don't have a centralised environment and blockchains are decentralised by construction. Therefore, you become somewhat vulnerable to selective denial of service attacks where I can try to slow down certain operations to avoid certain people participating in a certain way. There have been actual attacks of this kind.

In the cryptocurrency world the problem becomes more of a policing problem. If you have a completely untraceable way of due payments this could look very good, but could also bring a lot of criminal activity that cannot be blocked anymore.

Given blockchain's decentralised nature, can you talk about decentralised storage and its advantages?

Decentralisation has several advantages, which is fundamentally the ability to sustain an attack to some of its parts. Sometimes blockchain has been touted as a solution to a lot of problems. I think it's a solution to certain problems, but blockchaining everything is a mistake, it's a little bit of a fad. In order to verify certain transactions, you actually have to do a lot of CPU computation, that might not scale. The infrastructure might scale, but the proof of work that you have to do in doing a blockchain might require too much information.

So, imagine, for example, if all of us started using bitcoin. That means that every single transaction in the world will go through a ledger and has to be verified. Now suddenly there's so many of these transactions, or groups of transactions in the blockchain, that these blocks that have to be verified would need so much CPU power that we might not be able to provide enough electricity for it.

So blockchain is good for certain things. When you want to have a community that want to create consensus on a specific set of actions, for example. However, I would warn against using blockchain to solve every problem because it's not the solution, it's one solution; it's one building block to build in secure systems but not the solution to everything.

There are already cases in which governments have decided to prohibit certain forms of bitcoin mining because it was too expensive, and it was putting too much stress on the electric grid.

What are your top three malware nightmares currently?

I would say the use of machine learning on stolen personal data on an enormous scale. The ability, with all these breaches collecting data, is the possibility that somebody will really collect a ton of data and find ways to exploit this data to understand something about people; to understand who to attack or how to sell certain services based on this aggregate information.

Once this happens it's very difficult to make it stop because once that information is out, the 'cat is out of the bag'. Putting this information back might be impossible, so these large-scale breaches are worrying me.

I would say the second issue is this use of machine learning that is just wrong, learning the wrong things or learning things in a way that is too brittle. It can be bypassed, and I'm afraid it will bite us back in two to three years from now.

The third thing is evasive malware; malware that is able to easily escape detection or can escape detection from basic detection systems. Unless you get good malware protection it's really difficult to identify these evasive pieces of malware.

What's exciting you about cyber security?

I think that there are several things that are exciting. I am personally very excited about firmware analysis. So, firmware, in the internet of things, is this piece of code that runs very close to some hardware. The difference in analysing this firmware is that you don't have all the abstractions that a normal computer or even a phone has - the operating system and things like that, which makes the analysis a lot more challenging.

Usually you don't have the source code, so you have to operate at the binary level, which is more challenging. So, finding problems in firmware is something that requires new approaches, new skills and, for me, that's very exciting.

Find out about some security training available from TSG Training.