Privacy in the Blockchain Era

featuring Dr. Giulia Fanti as interviewed by Dr. Eliezer Kanal

Welcome to the SEI Podcast Series, a production of the Carnegie Mellon University Software Engineering Institute. The SEI is a federally funded research and development center sponsored by the US Department of Defense. A transcript of today's podcast is posted on the SEI website at sei.cmu.edu/podcasts.

Eliezer Kanal: Hi, my name is Eliezer Kanal. I am the technical manager here at the Data Science Group within CERT at the Software Engineering Institute. I am pleased to have with me today Dr. Giulia Fanti, assistant professor in the Electrical and Computer Engineering Group at Carnegie Mellon University.

Giulia Fanti: Thanks for having me. It's great to be here.

Eliezer: Let's start by having you tell me little bit about yourself. What you do here at CMU, and kind of how did you get here?

Giulia: I've been here about two years by way of a postdoc at the University of Illinois, and before that, a PhD at UC Berkeley [University of California Berkeley]. I work on a couple of areas related to privacy and blockchains and some machine learning related problems.

Eliezer: What kind of stuff did you do in your PhD work?

Giulia: During my PhD, I was working mostly on problems related to data privacy and anonymity, but, at the time, I was focusing mainly on social networks. Over time, I realize that people don't actually care that much about privacy in social networks and started reading more about cryptocurrencies and realized that this is an area that has both distributed systems, which is what I was focusing on at the time. It is a distributed system where people really care about privacy because these are financial transactions. I started gradually transitioning and applying some of the ideas that I worked on during my PhD to some privacy problems in the cryptocurrency and blockchain space. That is how I started moving into that area.

Eliezer: That is very cool. I want to get back to something that you mentioned about your PhD work for a second. You say people don't care so much about privacy, and then people do care about privacy in the context of cryptocurrency. Is it that people don't care so much about privacy, or they are simply not aware that privacy is actually a currency, and that it is a market that they are participating in?

Giulia: I think maybe a little bit of both. You hear a lot of people saying things like, *I explicitly do not care what Facebook knows. I have nothing to hide*. Privacy is one of those strange areas where you don't really realize the ill effects of losing it, because it is a lot more subtle than things like losing your credit card information. If you get hacked and someone uses your credit card to buy thousands of dollars of jewelry, you are going to notice. But if your privacy gets breached, it's a much more subtle thing, I think.

Eliezer: Is there an actual cost to people for having their privacy breached that we have been able to quantify?

Giulia: I think it depends. There have been studies showing that when people are in environments where they have lost their privacy—where they are being surveilled, and they know they're being surveilled—there are chilling effects. They do behave differently. But again, this is a much less explicit thing than having money stolen or having physical material possessions that are being taken away from you.

Eliezer: That is very interesting. I assume that these kind of chilling behavioral effects as you have mentioned, people will see it, and they may, if it's strong enough, notice it. But it is, for the most part as you said, pretty subtle. They are not going to feel it day-to-day as if someone grabbed their credit card or something.

Giulia: Yes, I think so. There are niches of society where people really do have to care about privacy like activists or journalists. But for the average person, I think it is a much more subtle effect.

Eliezer: That is interesting. This is I guess my own personal interest, have the fields given up on trying to educate that privacy has value, or is that something which is not even universally agreed upon?

Giulia: I don't think that people have necessarily given up on trying to convey that privacy might be an important thing. There are still studies that are definitely trying to understand how people interact with many of these platforms and also trying to understand what value they put on their privacy. There has been some interesting work at CMU on the economics of privacy. That is not my own area, but I think it is really getting at, first of all, trying to quantify the value that people put on privacy. Then, once that is quantified, then you can start to ask, *Well, how do we preserve it? How do we allow people to have agency over their own data*?

Eliezer: Very interesting. I know that for myself, I have had a hard time—I consider myself somewhat of an expert relative to the public in the field of privacy and what you can do with it, and I know that I have had a very hard time convincing any one of my friends, family, people I know outside of the research field that they should take steps to preserve their privacy. Also, in the same way, privacy is not binary. It is not like you have privacy or you don't. There always things you can do to have more privacy, and that usually comes at a cost of convenience. Trying to give that sense of it's a give-and-take—and is not just a one off or on—is difficult.

Giulia: Yes, totally. Convenience is a powerful trade-off, I think.

Eliezer: Yes, very much so. So let's use that, actually, to spring right back into the cryptocurrency. I actually teach a course here at CMU on it, and I am somewhat familiar with some of the research going on here. I think the way you mentioned it is, it's taken people even a little while to appreciate that cryptocurrencies have privacy vulnerabilities. I know there was some work here that kind of showed the extent to which we can data mine on the Bitcoin blockchain to identify who people are. Is that an area where you play around in?

Giulia: Yes, I've actually been focusing more at the networking layer. A lot of the work on showing privacy vulnerabilities in cryptocurrencies has been taking the blockchain and trying to do some data analysis. Linking together different transactions and trying to figure out who was doing what. That is one layer at which there are privacy vulnerabilities.

Another layer is underneath that, the peer-to-peer network on which it's actually running can be used to link transactions to the IP address of the person doing those transactions. That is another way that people can lose their privacy if they are doing transactions on a cryptocurrency network.

Eliezer: To step back for a second. Let's say I'm not a miner. I'm just someone participating, and I want to send money to colleague, right? I log onto the network, so to speak, so that I have a way to send a transaction. What you are saying is that network connection itself—put aside any transactions I may send—the fact that I'm connecting to the network is itself a vulnerability. Is that what I am hearing?

Giulia: If you try to generate transactions and send them over those connections, that can be a vulnerability. That can be used to link your IP address to the transaction that you are sending.

Eliezer: Does this tie in at all to Tor and Onion routers and such. Does that help protect in this context?

Giulia: It does, yes. Using Tor does help to protect against these kinds of attacks. Some of our research has been on trying to design different routing mechanisms and protocols that make these kinds of linkages difficult. We have a system called Dandelion that changes the way the transactions are routed to make this kind of linkage difficult, but you could also just use Tor, for example.

Eliezer: Is there any way you could convey how the current routing protocol works that would make sense to someone without a visual?

Giulia: Yes, sure. Right now if you think of the network as a graph of users, if I send a transaction, I will send it to my peers on this network, and they will forward it to their peers, and so on. This transaction is basically propagating like a ripple in a pond. This makes it easy to identify the source because you just have to find the center of that ripple to identify who was the source node.

Eliezer: Just to clarify, when you say peer, how do we become peers?

Giulia: This is done in a distributed fashion. So each node keeps a list of other members of the peer-to-peer network and chooses some subset of those—sometimes at random, sometimes not—to connect to.

Eliezer: The list does stay pretty much the same over time.

Giulia: It is roughly constant, yes. It is changing a little bit because people are coming and leaving the network, but it is not changing super quickly.

Eliezer: OK. As I kind of spread and ripple out, I guess I can start to identify how connections are being formed. Is that how this works, or how I'm getting from point A to point B?

Giulia: Let's suppose that you have a graph, and one of the nodes is originating a transaction. The graph, we can assume, is known to the adversary. There are ways to recover the structure of this graph...

Eliezer: ...and to translate that into non-graph speak, I know who can talk to who else.

Giulia: You know who can talk to who else.

Eliezer: And how they are connected, because I have seen it through these ripples before.

Giulia: In part through that, yes. In part, through other mechanisms.

Eliezer: Or other mechanisms, sure.

Giulia: So you have some way of figuring out who is connected to whom on this graph. If one of these nodes starts propagating a transaction, let's say like I decide to send you some money. The adversary, in many cases, can see that that's happening as it's happening. It can see that this ripple is reaching certain parts of the graph at this time. Five minutes later, it reaches this part of the graph, and so forth. Because it knows the structure of the graph—so it knows who is connected to whom—and it can observe some partial timestamps, some metadata on how this transaction is spreading, it is actually possible to do some inference attacks where the adversary can point to the source node.

Eliezer: In doing so, you call it an attack because the identity of the sender is supposed to be private, but we can infer it from, as you said, the metadata that is visible to the observer, to the attacker.

Giulia: Yes, exactly.

Eliezer: Knowing the little bit that I know about blockchain, so, why is the identity assumed to be private? Wouldn't the original message have the IP address attached to it?

Giulia: Good question. When you generate a cryptocurrency transaction, it typically has like a pseudonym, basically. It does have some identity attached to it, but it would not necessarily have your IP address attached to it. Users have different types of identities in cryptocurrency networks. One is the pseudonym, which is basically like their public key. One is their IP address, which is at the network layer. Typically, the second identity is considered separate and is not included in any of the messages that get propagated, but by doing these kinds of inference attacks that I mentioned, you could actually link these two identities.

Eliezer: That's interesting. To finish the whole story, so once I can identify person, and then I can identify the transactions coming from that person, all of a sudden, I get to this state where I can say, *Oh, this person makes these types of transactions*. So, to be fair, that's the current state of the world in credit card transactions. Is that an accurate assessment? Because credit cards, Visa knows very well what I purchased, because they process it. They see the originator of…Target, I just bought a banana. Maybe they will see the banana or not, but they'll know I shopped at Target, and they'll know my name.

Giulia: Yes and no. It is true that Visa now knows everything that you are doing, but what is particularly scary about cryptocurrencies is that anyone can do this attack. It's not just one party that you have already chosen to trust that can do this attack. You or I could go out tomorrow and set up a server to listen to the traffic on this network and try to do these kinds of deanonymization attacks.

Eliezer: So any untrusted party as opposed to one trusted party.

Giulia: Exactly, yes.

Eliezer: You have these fixes which you propose to this at the network level. We are going to try to see if we can extend that. I guess the goal is have a network, which is more private, so that you can't make this kind of an inference attack. You mentioned before, Dandelion. Is that something can you explain, how that would work?

Giulia: Yes, sure. Remember I said earlier that today, transactions are propagating like ripples. So, it's very symmetric. You're sending to all of your peers on the network at roughly the same rate. The idea behind Dandelion is very simple. Instead of sending to all of your neighbors, you are going to send to just one of them, and that one neighbor is going to send to one of their neighbors, and this happens for a few random hops. At some point, you transition to what we call the fluff phase, like the fluff of the dandelion. Then you start spreading symmetrically. It's a very simple protocol, but the advantage is that we can show theoretically that Dandelion achieves within very close to optimal privacy guarantees under some random model of spreading.

Eliezer: Why can't the adversary observe those first few hops?

Giulia: In the analysis that we have done we were assuming that the adversary has some number of nodes in the controls, and it places them randomly in the network. The kinds of guarantees that Dandelion gives is average-level guarantees across the whole network. If an adversary is particularly interested in you, they could certainly monitor your neighboring edges. Then, this wouldn't work.

Eliezer: That is interesting. So, the threat model that we are building here is against an adversary who is looking for essentially anything, not a targeted attack?

Giulia: That's correct, yes. There are a few companies today that are emerging whose entire business model is to do blockchain analytics. They are basically trying to deanonymize users. This is more targeted towards these kinds of large-scale surveillance efforts, and less towards targeted attacks.

Eliezer: As someone who has followed blockchain technology broadly for a little while, I'm guessing you are referring to the insecure protocol, I'm guessing you're referring to gossip, is that accurate? The gossip protocol?

Giulia: Yes. What's done today is one example of a gossip protocol.

Eliezer: OK, something similar to that, sure. I know that there are a whole lot of blockchain platforms out there, in fact, more than I could probably reasonably state in the entire allotted time on this podcast. Are there certain platforms which have completely different protocols for which this is not a concern or for which even if we implement the safer approach, it's still a concern? Are you aware of any like, *Oh, this is a good platform*.

Giulia: I know that there are some platforms that have different networking stacks. I am not sure offhand. I wouldn't be able to do this comparison offhand of whether some of them inherently get around this problem. One thing to notice though is that the gossip protocol that we talked about earlier is used in Bitcoin Core, and a lot of other blockchains have forked the code of Bitcoin Core. The networking stack has kind of been ignored for most of the 10 years that blockchains have been around. So, this particular aspect of blockchains is present in a lot of projects, not just Bitcoin.

Eliezer: Well, one thing which I've always found interesting is that encryption is not a solution to this. Many people think, *We'll just encrypt the data*, but in that case, example I tend to give is, *You're just sealing your envelope better, but the address is still on the front*.

Giulia: That is right. Yes.

Eliezer: Even if you encrypt the packet, I still have to observe how it gets to wherever it's going. I guess for your purposes, it's not even something you'd consider, because you're looking, as you said, at the network layer.

Giulia: That's right. Yes. So encryption doesn't really solve this problem, because at the end of the day, you're trying to reach everybody.

Eliezer: Let's take a total and complete left turn here. I saw, looking at your faculty page on the CMU website, I noticed you also have some interest in generative adversarial networks, or GAN. Is that something which is related to this, or is that a totally separate line of research?

Giulia: It's pretty separate at the moment, but also interesting, I think.

Eliezer: That is fascinating stuff. Do you mind going into little bit about what you're doing there?

Giulia: Not at all. So, maybe I'll start by explaining what a GAN is. GAN is what's called generative models, and the idea is the following. Let's suppose we give a GAN a bunch of samples from a particular distribution. Let's say we give it a bunch of images of faces. What the GAN does is to learn from those faces how to generate new samples of faces that are not just copies of the samples that I gave you earlier, or that I gave GAN earlier. It's able to draw random samples from some underlying data distribution.

People are excited about GANs, in particular, because they have been able to generate some of the most photorealistic images of any generative model for the last decades. Generative modeling has been a problem with interest for a very long time. These are also the technology behind deepfakes, which you probably heard of. People are using GANs, in some cases, for nefarious purposes, to generate videos or images that are trying to depict people doing things that they didn't actually do or saying things that they didn't actually say. That's concerning, for sure.

On my end, I'm interested in using them to generate privacy-preserving data sets. The problem there is that a lot of companies have data that they would like to release to researchers to either develop new algorithms or like test out new ideas, but they can't release this data for privacy concerns, and this is true even within the same company. Sometimes one department collected some data, and another department would like to use it, but they can't because of privacy restrictions.

We are trying to use GANs to generate synthetic models of data, specifically, networking data, time series, data traces that resemble the original data. They look very much like the original data, but don't have these privacy concerns, because they're not associated with any individual's information.

Eliezer: To be fair, also you mentioned before, this is networked data. So you can use GANs to make deepfakes. You can use GANs to do cool video and audio work. You don't have to. They can be used for any type of data, as you mentioned, network data, packet flow, or whatever it is you're working with.

Giulia: Most technologies that are interesting can be abused or used for good purposes.

Eliezer: To that extent, when you talk about something being private, broadly speaking, I guess it has to do with the individual data point, right? I know that person's Social Security number. It's not a distribution. I guess the goal here then is to mimic the entire distribution. Are there cases where the distribution itself is a privacy concern?

Giulia: Yes, great question. In some cases, for example, if you have some very sensitive data points in your training data that you give to the GAN, that could be reflected in the model that the GAN learns. One of the big questions is how do we train models that aren't leaking information about may be the most vulnerable or the most unique elements of the training data set?

Eliezer: One thing which we've been concerned with here at the Software Engineering Institute is a technique called model inversion, where you can look at a completed, essentially black box model. There's a model in a box, and you can ask it questions and get answers. Though that you can actually identify data points in the original training set. You can imagine that if the original training set contained healthcare information or the names and faces of people who don't want to be public, that can be a negative thing. Do you actually see the ability to perform model inversion on GAN-generated data sets?

Giulia: There are definitely similarities in the samples that GANs generate and the original training data. Right now one of the main areas that we're looking at is trying to understand the privacy guarantees that GANs give you and how to generate privacy-preserving GAN models. We have been doing some experiments on membership inference, the kinds in the text that you're mentioning, where we try to infer whether a particular individual's data was present in the training data or not, based on just observing the output model. It seems like one trend that appears is the more data you train on, the less effective these kinds of attacks are, which kind of makes sense. You're averaging over more data points. We're not at the point where we can make these kind of claims rigorous, but the goal I think is to be able to say, *If you train on this many people, and your distribution has these kinds of properties, then you get these kinds of privacy guarantees*. That's the goal. That's where we're trying to work towards.

Eliezer: There's a technique that has made a little bit of news recently called differential privacy, where instead of storing it at full data points in a database, actually storing info about a lot of points in the database. It seems like the approach that you're doing is, *I'm going to create a data set that looks like the original but isn't*. And that approach is, *I'm going to store the information about the data set rather than the data set itself*. Am I reading that correctly? Are those two similar, or is that actually different lines of research?

Giulia: You can actually incorporate differential privacy into machine learning models. There is a whole line of research on people trying to train machine learning models in a differentially private way by adding noise during the training process. If without privacy you would've updated your parameters by epsilon. Under the differentially private version, you updated by epsilon plus some noise, just to paint with a very broad brush.

There has been a lot of work on this on differentially private gradient updates for machine learning. The problem is, it adds a tremendous amount of noise. The final models that you get tend to be not as useful as one would like. That is particularly true in GANs, because you have to do so many updates during the training process that the amount of noise that you end up adding is really substantial.

Eliezer: This is interesting. In one context, I will have a very private but very noisy data set. In another case, I may have some more privacy concerns—we are still researching that it sounds like—but the data set will have much higher fidelity. I guess the deciding factor for someone trying to adopt these things is really, *What is their level of risk that they require for their application?*

Giulia: That's right. Yes, I think there's definitely a tradeoff. People have actually tried training differentially private GANs. It is an issue of fidelity versus privacy there. If you want reasonable privacy guarantees, you end up getting a very good fidelity and vice versa.

Eliezer: Let's take a step back for a second. The Software Engineering Institute, as many of the folks who listen to this podcast probably know, is a federally funded research and development center. We are funded by the Department of Defense. We do all sorts of nifty work that helps the folks out there who are on the ground accomplish missions that they're trying to accomplish. I believe though you are in the Electrical and Computer Engineering Department, but under the broader CyLab umbrella. Could you talk to us a little bit about what CyLab is and what they do here at CMU?

Giulia: CyLab is CMU's privacy and security umbrella organization for people working on research or education in that area. It takes a pretty broad view of what constitutes security and privacy work. It includes people from all kinds of different departments, including my own, ECE, mechanical engineering, computer science, public policy. It's a very, I think, inclusive organization and lots of collaborative work going on, in terms of security research and education as well.

Eliezer: Within CyLab—it sounds like there's a lot of people in the different departments. Do you collaborate with those people? Is there a mechanism to help people who are doing similar research talk to each other?

Giulia: CyLab does have its own space in the Collaborative Innovation Center. A lot of us from different departments are in that space, so that makes it a lot easier to talk to people from different areas and organically start collaborations. There is also a bunch of people who don't sit in that space but are also affiliated with CyLab. There are periodic events and small conferences where people get together to discuss potential research opportunities and collaborations.

Eliezer: Thank you, Dr. Fanti, for joining us today. Thank you all for joining us, as well. We will include links that reference the stuff referenced in today's discussion in the links below, and you can find more about us at sei.cmu.edu. Thank you very much.