Incentives in Security Protocols (Transcript of Discussion)

Steven J. Murdoch

University College London
s.murdoch@ucl.ac.uk

Steven Murdoch: Whenever there is a risk of a security problem, the person who caused that security problem to take place should be the one who actually takes the risk. If the merchant skips PIN verification, the merchant takes more of the risk, and if the card-scheme does stand-in authorization, the card scheme takes more of the risk. In this way, they try to encourage everyone to move to a more secure system. There are problems with that; this is called liability shifting. If one party is able to move much faster than the others then you have the issue that suddenly risk gets dumped on one party, who is not necessarily acting any less securely, but is just acting less quickly.

Daniel Weitzner: Can you give an example of liability shifting in the EMV case?

Steven Murdoch: The particular problematic case is right at the start of the roll out of EMV. The rule was that if a terminal is capable of a chip transaction and the card is a magnetic stripe card, the bank pays the fraud, but if the terminal is only magnetic stripe capable and the card has a chip, then the merchant pays the fraud. But it turns out the banks were able to very rapidly roll out semi-functional chips that weren't really useful (they were very slow and buggy). However, it was enough to trigger this liability shift. The merchants were much slower about rolling out their terminals because a terminal would last three years or more, whereas a card would only last one or two years, and suddenly the merchants had a huge amount of fraud that they were then having to cover even though this wasn't really their fault. Even more problematic is when liability is being shifted onto the customer, because when liability gets shifted to the bank they find some contractual way to dump it onto the customer, even though the customer is not in a position to make things more secure. This sort of liability engineering can be effective, but it is only going to be effective if there is actually sufficient evidence in the system that allows the liability to be fairly assigned to the right party, and in EMV that is generally not the case.

There are logs, but these are debugging logs for developers and there is a number of problems with that. The first is that if there is any disagreement between different aspects of the system, say because of an attack or failure, one side sees one aspect of a transaction. Say that the PIN was verified correctly and the other side sees that the PIN verification was skipped, debugging logs will only tell you one side of that story. Because the debugging logs are not actually parts of the functional requirements of the system, often it is not written down what aspects of the system they are showing. The second thing is that these debugging logs are not very good for presenting in court, even though ultimately a jury or a judge is going to have to interpret them. When I was an expert witness in one of these cases, it was simply the hexadecimal code ten and then the bank expert said that this shows that the customer is liable, and it just says ten [laughs]. There was no documentation explaining why ten is actually an explanation for this, it was fairly unconvincing but that was the only evidence that was available. Then yet another issue is that this evidence can be tampered with; it is just stored on the developer's machine for whatever reason that they need. It is not going to go through the same chain of custody that you would expect for evidence that is actually deciding hundreds of thousands of pounds of money being transferred from one party to another. It is not actually a complete disaster but this is more through historical accident than by design.

When the banking payment system was set up, cryptography was expensive. You had to use expensive and slow line encryptors. You needed to deal with all the key management and so rather than doing end to end encryption, which is what you would do for say, Internet payments, the way it works is that communications are on a point to point basis and each side of this communication has some sort of contractual agreement. The customer talks to the merchant and they have an implicit contract. The merchant talks to the acquiring bank and they have a contract. The acquiring bank talks to the card scheme and they have a contract. The card scheme talks to the issuing bank and they have a contract and then the issuer deals with the customer again, and then they have a contract. Because these legal contracts are set up, you can sort of get away with not doing encryption because you do have some assurance that people are not acting completely maliciously.

Daniel Weitzner: I feel like there is something behind this claim that this was a historical accident. I have a hard time accepting that at face value. I can understand your point that there were some technical constraints, which shaped the way the parties aligned their liabilities but I am not sure why you think that is an accident. That just seems like an adaptation and so you are saying it would be better if the liabilities were aligned differently or if there were more options available?

Steven Murdoch: Okay, I will clarify what I think I mean, which is that the fact that the evidence is somewhat reliable is a historical accident that comes from a particular architecture, which is designed for good reasons, which were valid at the time and are still valid. If however, there was no such arrangement and instead communications just went over the Internet, straight from the customer's phone to the merchant's phone, did not involve intermediaries and had end-to-end encryption to deal with security, then the quality of the evidence would not be as good because everything would be encrypted. There would be no clear-text logs that someone can show, so in the case of the number ten saying the customer is liable, at least someone who wasn't a party to the dispute had that number then. If it was end to end encrypted, even that would not exist.

Daniel Weitzner: It sounds like you think that is a bug. It sounds like a feature. I am just not sure what we are supposed to conclude from the fact of this accident. Is it that you couldn't use these in any legal scheme, which seems obviously to be the case. I am not sure what your solution is for this.

Steven Murdoch: It is a bug and a feature. It is a feature because at least there is some evidence that you can show. It is a bug in the sense that the design of the system for creating evidence was not very well thought through, so the evidence is somewhat poor. So we're somewhere in the middle, we have got evidence and it is sort of okay most of the time but not really. If the system was designed in a different way, there would be a realisation that there is no evidence whatsoever so we need to build an evidence overlay and then reach the other extreme where actually the evidence is very good. We're in this sort of middle where it is not great and it is not a disaster and I think that is the historical accident.

Ian Goldberg: You said the customer has a contractual relationship with the issuing bank, the merchant has a contractual relationship with the acquiring bank and the banks have relationships with the card scheme. The latter one I buy, the banks certainly have contractual relationships with the card scheme. The relationship between the merchant and the acquiring bank somewhat so, but the relationship between the customer and the issuing bank really is a contract of adhesion. There is no negotiating this contract on the customer side. The customer is just presented with "you want to use this card, here's your 10 pages of terms of service that you have to agree to because if not, you do not get a card", and certainly when they rolled out the chip cards in Canada, you basically had no choice. They said when your card expires the next one you get will have a chip and these are the terms that come with it and it involves all the liability shifts from the bank or merchant to the customer, but the customer of course had no say in this. These contractual relationships that you might want to lean on to decide where the liability goes, maybe morally shouldn't even be considered that because the customer has no say in it. If the contract shifts liability onto them, it is really not their fault.

Steven Murdoch: Yes, the situation is not good. I've acted as an expert witness in both dealing with customer cases and merchant cases and you are right, there is no real negotiation here.

Daniel Weitzner: But hold on, your complaint is with the bank regulators, because it is the bank regulators who ultimately either affirmatively consent to whatever these new terms are or just aren't. Or they are asleep at the switch, intentionally or otherwise.

Ian Goldberg: Captured regulators?

Daniel Weitzner: No, no, but seriously you look around the world and there are very different arrangements. Specifically about consumer liability, U.S. law has a couple of different rules and they make these judgments. I think you can dig down deeper and there are some contractual terms that maybe the regulators do not initially need to pay attention to, but it is still not quite a fair bargain. But I think as to these broad liabilities for failures of whatever sort, that is squarely up to regulators. The question is whether they are doing their job or not.

Steven Murdoch: Yes, so from my perspective the U.S. regulators are mostly doing their job well. But that really came from Jimmy Carter, when he was president. He set down some rules and those are still the rules, which roughly say that the customer is never liable. There is some 50 dollars, 150 dollars, but in most cases that is waived. The UK is somewhere in between because there is a Payment Services Directive, which is written moderately well with a couple of bugs, enforced quite badly in the UK, better elsewhere, and Canada is actually the worst in my experience, but it is because the regulator basically said the banks can do what they want.

Okay, so that was one example, EMV. Another one is on cryptocurrencies, we have all heard of cryptocurrencies. These are distributed, decentralised to a certain extent, and because you do not actually have any contractual relationship between anyone really, it is purely functioning on the basis of incentives. That makes it a little bit fragile, because if the incentives are not aligned properly then you have problems. The protocols that are designed for cryptocurrencies are often reasoned about in the ways that we reason about security protocols nowadays. We use formal models, we use proving techniques and model checkers and all these sorts of things. But when it comes to reasoning about incentives, it is sort of like being back in the 1980's for protocol design where, if someone proposes a protocol, they think that they are not able to break it, they show it to a room like this and nobody is able to break it, then it is good to go and they ship it. They start putting billions of dollars through the thing and this is fragile, this is problematic and you do actually have failures. For example, one set of failures, for example selfish mining, come from assumptions that Nash equilibria are the right way to think about incentives. Nash equilibria make a whole bunch of assumptions to do with parties being asynchronous and parties actually acting rationally and then when these assumptions go wrong you have attacks that are going to be possible. The other place that incentives have a role to play are to do in the fail-safe and the fail-deadly aspects. An example of a fail-safe incentive model is, you need to make some change to the software for whatever reason but you can do this in a backwards compatible way, so that the people who are possibly in a failure state because they are running the buggy version of the software are still able to interoperate with the other clients, which are running the more correct version of the software.

Ian Goldberg: Okay so this is something I do not get about soft forks in blockchain type things. This may be a little tangential but maybe by explaining what exactly you mean by this, this will clear this up for me. There was just a talk at Financial Cryptography a couple weeks ago, where they talked about soft forks, hard forks, velvet forks and I do not know, tiramisu forks or something. But when you do this fork, the set of valid transactions changes between the old software and the new software. In a soft fork, the old software will still produce transactions and blocks recognised by the new software, but it might reject blocks produced by the new software. As soon as the first block is mined by an upgraded client, what happens to the old client? The old clients, they are still mining. They see a block and they are like, “that is not valid, throw it away", and they are still mining on the old thing and now you have an actual fork in the chain. How is that fail-safe?

Steven Murdoch: I think that is the intention of the fail-safe, the reality may not be. I do not know if Sarah or Paddy or anyone wants to say something.

Patrick McCorry: I can say something about that. The whole point of the soft fork is that the miners are enforcing the new rules and I am not convinced it is actually incentive aligned. One of the incentives of the network is that you can verify everything yourself, but what you are doing in the soft fork is that when you create this new block, with the new set of consensus rules, you did it in such a way where you trick the old clients. All they see is an empty transaction was sent, they cannot validate the rules at all. You rely on the miners validating the rules and over 51% of the network enforcing it. Only operating clients can see the new rules, that is the soft fork. It relies on the fact that the miners are enforcing the new rules and you can trick old clients, and the fact is that there is some trickery there, maybe your comment earlier with the decisions and you do a magic trick, a sleight of hand, I do not know if the incentives are fully aligned for that.

Sarah Azouvi: Did that answer your question?

Ian Goldberg: Sure.

Steven Murdoch: Yes, that is the intention behind the fail safe even though it might not work, and then the sort of fail deadly approach is you have a chain split. For example, if Ethereum is splitting from the Ethereum Classic. Value does get destroyed for some people and you would hope that the incentives are aligned such so that people who suffer are the ones who have the money taken off them. Although, I am not actually convinced that going to be the case there.

Mansoor Ahmed: I am just wondering if it is even possible to design an incentive compatible system where there could be a theoretically infinite number of nodes. For example, many alt-coins claim to have an incentive compatible system, but then we see alt-coin infanticide where Bitcoin miners just decide to kill that alt-coin even though it is not incentive compatible. In a system where anyone can join, is it even possible to have some sort of a consistent incentive structure?

Steven Murdoch: That is a good question, I do not know. Paddy do you have a thought on that?

Patrick McCorry: I have one more comment, it is sort of like a tragedy of the commons. Joe Bonneau highlighted that, all of the miners have a long term interest in the health of the ecosystem and the blockchain itself. But in the short term, if they can boost their short term profits i.e., killing an alt-coin or mining an alternative cryptocurrency because they get more money, we have seen that with Bitcoin and Bitcoin cash, they are actually going to do that.

Alexander Hicks: One thing with something like incentive compatibility, usually you read the paper and they say their protocol is incentive compatible, but that is for incentives in the protocol, which might not relate to people outside the protocol that then decide to kill the coin, which is the problem you have a lot, the tragedy of the commons. People are going to compete to make their coin the most valuable. We'll get to incentive design later on, but it is worth taking time to point out that when people talk about incentives, they usually take into account only incentives in the protocol they designed rather than incentives for all, which is where a lot of problems arise.

Mansoor Ahmed: But if you have an open membership list, does it even make sense to talk about incentives within your protocol without looking at incentives for the whole world.

Alexander Hicks: Yes exactly, that is a failure of the models that are used now.

Steven Murdoch: It is actually linked to the next and final example, which is Tor. For those of you that do not know Tor, this is an anonymity technology. You send your traffic via three intermediaries and then your traffic comes out through the other end in such a way that it cannot be traced back to where it came from. The people who run these servers are volunteers, they do not get paid for it. A handful get some bandwidth reimbursement through a government scheme but most do not. There are 6 000 out there who are not actually getting any economic benefit from it. Looking from the fail-safe and fail-deadly aspect, the fail safe approach is that some of these may be malicious but there are three of them so unless the first node and the last node are colluding or being observed, you should still be safe. The idea behind fail-deadly would be that if a node is detected to be misbehaving, and there is active scanning for this. Then you could kick them out and then prevent them coming back. But there is Mansoor's point about what happens when there is an indefinite amount of people in the system. Also, what happens if you are not able to identify the people operating these servers? The person will probably just come back again and they will be throttled for a while, but eventually they will go back to the previous state, under a new identity. The other interesting aspect about Tor is that there are concerns about introducing monetary incentives. The example that is sometimes given is that there was an economic experiment done in an Israeli nursery scheme where the problem they were trying to address is that parents were coming late to pick up their children, so they introduced a fine if you came late to pick up your children. It turned out that parents started coming later because you'd moved a social punishment of just feeling bad and maybe getting into trouble, into an economic punishment and they just considered this is the price of extra childcare and they are very happy to pay that fine as a price for extra childcare. At the end of the experiment, the nurseries removed this fine but still the parents came late. By shifting from a non-economic to economic incentive scheme, you've actually permanently damaged the system and that is why things are not changed.

Mansoor Ahmed: I understand why you would not want to do monetary incentives but are there trepidation incentives. Is there a leader-board where we can say "oh I contributed this much bandwidth"?

Steven Murdoch: I've got a Master's project on exactly that so maybe we should talk about that. Currently there is a leader-board but it is fairly simplistic. So the idea behind this Master's project is to actually take a little bit of psychology, game design and marketing and then use this to make it a bit more fun to run a Tor node.

We have already had some good discussion, here are some other points that we could consider to get things started, I do not know how much time we have left. How do you reason about security protocols from the incentives perspective? How do you choose the right kind of incentive? There is a categorization in the paper where we can look at different incentive schemes and different enforcement mechanisms How do you actually enforce them? Do you need to have a regulator? Do you need to have strong evidence? How do you actually do that and then should you use things like Nash equilibria, there is the BAR model, rational cryptography. When do you use a particular model in the particular context?

Ian Goldberg: I actually have a question about a figure in the paper, which did not appear in your talk. Right at the end of the paper, figure two in the pre-print [removed from final version], you have this Venn diagram here where you have different models by field. You have three circles, Game Theory, Cryptography and Systems, and you have things filled in a bunch of places. Notably the cryptography and systems intersection is completely empty. Why? Is there really no intersection between Cryptography and Systems?

Frank Stajano: Theory and practice people?

Steven Murdoch: What would you put in there?

Ian Goldberg: I mean, there are a lot of things that touch both Cryptography and Systems. Any real protocol design for example, will have both cryptographic aspects and systems aspects.

Virgil Gligor: I will cover one of those

Ian Goldberg: Sure, yes. Like when we built Off-The-Record, there were very specific design choices we had to make to make it both cryptographically correct, but also there was a maximum message size that some networks supported. We had to make sure that instead of sending this message in message one, we have to send a commitment to it in message one and then reveal it in message three and things like this. I think protocol design certainly sits in the intersection of cryptography and systems for example in many cases.

Sarah Azouvi: In this diagram, what we wanted to put are the formal models that people are using to reason about security. There are a lot of protocols that combines cryptography and distributed system, and blockchain is one of them, but what you see is, for example for blockchain, what they use in order to prove formal models is more from the cryptographic literature or more from the distributed systems literature. What we are saying is that maybe we need new formal models that can encompass this better, because these models have failed to encompass a lot of attacks.

Ian Goldberg: But what about things just like Dolev-Yao and pi-calculus. These things definitely look at both the cryptographic side and thee distributed systems side and model the actors and model their messages. I think these things would fit in this section here.

Virgil Gligor: Just a comment. What has happened here is that Steven drew a boundary, which is reasoning about these protocols. There is always some other mechanism below the boundary that of course is not addressed in here. Indeed there is an intersection between Cryptography and Systems but at a much lower level than these, so in that sense the diagram reflects this abstract level as opposed to more concrete systems level.