Simba - transcript of presentation video

Paul: So, good afternoon everyone, my name is Paul from Lloyds Banking Group and I'm going to be presenting with Kurt from Duality, and we are proud to represent Team Simba. Team Simba is a multi-disciplined team, made up of people from the not-for-profit Cyber Defense Alliance, from a number of banks, and also from a tech company called Duality, which is from the Team 8 stable which is based in Israel and also the US. And we want to talk about our pet, our pet is Duality and our pet is called Simba, and we would like to give you a demonstration and show you over the next 5minutes.

I think the key thing that we all recognise is that criminals are not loyal to particular banks, they do not do all of their transactions through a specific bank. In fact, they do the opposite, they transact across lots of banks, they don't discriminate, they move accounts, they close things, they move around, that makes it really difficult for banks like Lloyds Banking Group. In order to investigate criminality, in order to KYC, in order to know your criminal, and what we are doing with the platform, from Simba is really trying to provide a platform that is going to help banks like Lloyds, to know your criminal better, and to do it in a fashion, that means that we don't need to impact on the privacy, and on the sensitivity of the data that the banks hold.

What we want to do, is we want to illustrate this, using the data that the FCA provided at the beginning of the week, and for those who have seen the data, there is a number of banks in that data, there is lots of data there, some of you may have seen HCBG, so for this presentation, just assume that its Lloyds, let's think about that as Lloyds Banking Group.

We've done the analysis on that data, we have actually picked out a number of things that are quite suspicious, and this is one illustration of many that we found, and on this illustration here, I don't know if you can see it from the back, there is a number of dots, each dot is an account, the grey dots and the green dots are in the UK. The green ones are Lloyds, the grey ones are other banks and they are all gravitating transactions towards the central dot there, that you can see in the middle, which is the Lloyds account, and from that Lloyds account, the transactions are being dispersed across those red accounts there which are accounts internationally. So, in other words, the money is moving abroad, okay, so this looks suspicious.

So, to investigate this, of course, we began to look at that central account, it looks clean, there is no suspicious flags on that account whatsoever. So, we then start to look at the other Lloyds account that are transacting with that account, and what we found was that, 68% of those other accounts have got some kind of suspicious indicator of one type or another against them.

So, again that doesn't look good, so we want to investigate this further, we want to go to the next step. The next stage unfortunately, means we need to talk to other banks, and get information access to the data from those banks, the next account, and the next level and so on, and we can't just go and do that, not everyone will share sensitive information, so maybe we go to law enforcement. Have we got enough information for it to be on law enforcement agenda, they are quite busy, is it criminality? is it big enough? We are not sure. Even if they think it is, they've got to raise a lot of paperwork, they've got to issue it to all the banks, they've got to collect the data, it will take time to come into different places, they've got to analyse it, and that is going to take months.

Ok, so the next stage is going to take months, and for us it feels like we've reached a dead-end. It feels without a lot of persuasion, without a lot of paperwork, without a lot of extra effort, we are going to not go much further. What if we had a platform, technology platform that would allow banks like Lloyds to go to the next stage immediately? To begin to use that technology platform to analyse the data, and in effect build up this picture here. Now this picture on the network, we're talking about network of accounts, is the totality of all of the banks put together, but this is a picture that no one can see, its hidden in the data, it's the visual, it's the visualization that no one can visualize at this stage, and so what we want to do, is in-effect use the Simba platform, and it means all of the participants need to homomorphically encrypt in a pre-agreed set fields.

They homomorphically encrypt and put it where they want, they can keep it on perimeter if they want, they can put it in the cloud if they want, they can give it to a third party if they want, but its protected, no one can access that sensitive data, and then what we can do with Simba is ask some homomorphically protected questions against that data, that will not reveal the underlying data, but will give you some insights.

So, for example, on those accounts we see a common mobile number, how many accounts overall are using that mobile phone number, that gives you extra insight. We can do that on account, we can do that on transactions, and what Kurt is going to do now is to give demonstration of the Simba platform.

Kurt: Great, thank you Paul. So, I'm Kurt Rohloff, CTO and Co-Founder of Duality Technologies, and we are going to start with a demo, of the actual capability which addresses the need that Paul discussed, and particularly what we are doing is, we are starting with an investigator who has identified a suspicious account and using that suspicious account to generate a query that they want to run, except in a privacy preserving manner.

So particularly what was going on, what you are seeing here is generating a query with this account number, which is encrypted to generate an encrypted query, to suspicious transaction accounts and other banks and identify which of these transaction accounts are associated with flags indicative of suspicious behaviour. And one thing I want to stress is that, as we are putting forward on this, is that the underlying technology is homomorphic encryption and most importantly with this demo video, this demo is being run in real time. So, you can see the responsiveness of the technology, particularly in the support of homomorphically encrypted queries, for running computations on sensitive privacy and sensitive data.

In particular, running over multiple data sets, multiple banks data centres assimilated in the Amazon AWS environment and each of this banks has hundreds of thousands of entities across these multiple data centres and the goal of this, is to support investigators while protecting the banks' customer information, not exposing the subject of the investigation and protecting the privacy of the individuals and so we see right here is what we have returned are 4 accounts associated with 4 suspicious flags, which we can then use to run further computations on, in particular capability of applying a complex machine learning ago (notably XG Boost) to run over bank data, to classify the behaviour, in a privacy preserving manner as suspicious or not, to help indicate and direct the human centre behaviour further investigation.

So, moving forward for this, the big idea for this is, Data Science and Privacy at Scale and particularly to support investigations with data across multiple banks, cross-sector industries and agencies where private data resides only where the bank wants.

This technology is being pushed over many years particularly by my team and also investment that DARPA has made in this technology, where I have been running project for DARPA in order of ten years to develop open-source versions of these technologies which are integrated in team Simba to extend beyond the classical limitations which have been discussed earlier associated with homomorphic encryption. Critically to provide data science optimised to run on encrypted data for general computation, queries, and execution of complex models.

In particular, associated with this and one of the major lessons learned from us, from this TechSprints activity in the week that we spent with the investigators and AML Experts, is the fundamental market readiness of this technology for general capabilities ready to take in the proof of concept, particularly deployed on portable and quickly integrated standard Linux stacks. It integrates also standard Data Science frameworks and backends in a scalable, efficient and responsive manner. Paul.

Paul: Thank you Kurt, we are convinced that homomorphic encryption is now ready to be used, to be used in the real world, and to be used at scale and that we are going to press ahead with the proof of concept to prove that, and we think we can do that in a way that those address the privacy concerns that people quite cant rightly have over some of this information, but not only would it help Lloyds and it will help Lloyds, we'll be able to do things a lot quicker, we'll be able to weed out the really important investigations from the less important and so on quickly.

It should help everyone, and if it can help everyone, then hopefully we can begin to make this digital jungle a better place and address some of those criminality (unclear).

Ok, thank you everyone.