Watch Transcript-Monrose

Keith Marzullo: OK – Um, I'm happy to introduce our speaker for our 7th WATCH. Let me read his bio to you. This is Fabian Monrose. Fabian is an Associate Professor of Computer Science at University of North Carolina at Chapel Hill. Prior to joining UNC he was an associate professor at Johns Hopkins, and a founding member of the John Hopkins Information Security Institute. From 1999 to 2002 he was a member of the technical staff at Bell Labs Lucient Technologies. He's received several awards including the National Science Foundation CAREER Award 2006 – Always good, and best paper awards at flagship security conferences including IEEE Security and Privacy and USENIX Security Symposium. He got his Ph.D. and Masters from the Courant Institute of Mathematical Sciences at New York University.

Fabian Monrose: Thank you very much. It's great to be here. And see so many familiar faces, some long time friends…that's good to see. And today what I'm going to try to do is recap some of the work we've done over the last 5 years and try to take it from a higher level perspective that you would normally see at an academic conference talk in the sense that I'll try to give you the intuition behind the approach we took and some of the reasons why, I don't want to say attack, but the information disclosure we discovered are possible, ok?

So this talk is entitled Hooked on Phonics. The string symbols in the title will be explained to you a little bit later and the running title here is: Learning to Read Encrypted VOIP Conversations. I should say VOIP sessions. So, this seems a little odd at first, I'm going to try to read something that is encrypted and particularly I'm interested in VOIP conversations. So the first thing is …..(talking to tech help) so as we all know, VOIP is increasingly popular as a replacement for traditional telephony and in most cases I think probably everybody in the room has used some voice systems in the past, right? So I probably don't want to spend too much time there--But several analyst predict that the subscriber base that there are 25 million subscribers by 2013. And some of the reasons for this is obvious, cost is one as in a case in point this little known school just north of where I am, last month announced that they had completed their transition to VOIP, so the entire system has moved to Voice Over IP, and they believe they will save something like $2.7 million annually by this transition. So, of course not to be outdone, UNC said we'll do the same and I'm on the board to see what we are going to do. Alright? So--So there is the cost savings. VOIP you get all the benefits that you get with a traditional landlines, right? And in particular you also get some benefits that we might not necessarily have with landlines. So first, for example, you can get email transcriptions of your voicemail. Alright? And the subscriber really base took off when we turned the corner allowing you to use your existing phone number landline and use it as VOIP, right? So, we are all attached to our phone number. We really didn't want to give it up but now with VOIP you can now do that. You can have-- Make 911 calls with some of the systems. You can use VOIP on your mobile phone, on your laptop. We all use it. It's great. So, I think what is driving the subscriber adoption though from the client—From the user perspective is that we have many free online services. Right? So, we have Skype. I will get back to Skype in a second which is great and a lot of these services are remarkably reliable and relatively easy to use. Right? So, you have used Yahoo Messenger, Google Talk….and Skype. They are really relatively easy to use. My mom uses Skype so you know…My mom can use that we have no problems making even video conference calls. It works just great and you know as a case in point on the the fact that these things are reliable and relatively easy to use. I think Skype a few months ago Skype announced that they had something like 65 million user online every day. They have made more than 300 million video calls every day and something like 30 million Skype to landline to mobile phone calls everyday…from the security perspective though and privacy implications of VOIP is still not well understood…The majority of the tension and the security community really focused on the…So if you think of making a phone call at the VOIP level. You know if I have to call Farnam and I have to figure out where Farnam is so I might use something…Session initiation protocol to figure out…So you make these connection and it might be connections…what is called sigma proxy and try to locate where Farnam is and the case …prior to like Skype. This might be over peer to peer appear network…but once I have done the discovery then you have this connection between the two devices, okay? And from a privacy persepetive this is—You know this connection will be encrypted, right?...and used something like the real time transfer protocol or SRTP which is a secure version of the real time transfer protocol. So, all the attention in the last five maybe eight year really focused on caller id spoofing…service attack all focused the call setter phases. And in particular the lack of authenticity and integrity checks in the (SIP) protocol, okay? And which allows you to do things like what is now called (SIPT) which is spam over IP telephone, okay? We are interested in something different and so in particular to explain what is going on. Let's take another look at what's happened to VOIP…there are two channels. There is the voice channel and there's the control channel. And all the work on prior attacks is focused on the control channel…right um—we are really, really interested in the privacy of the VOIP voice channel, okay? So, after this communication is set up…I am interested in the quality of the voice channel…encrypted. So you have an expectation of privacy, right? At least I do. Right? And um—So that's where we are focusing on. So for us, we are looking the accessibility of eavesdropping and encrypted calls after they have been legitimately established. And by that I don't mean there's this…person in the middle—Man in the middle attack, right? You have a legitimate call set up, you know we are trying to figure out what can you do then, okay? Now so I have to give credit to ….and Susan Landau for their book club. This image comes from privacy online which focuses on—Which I think is an excellent title for this work. They focus on the conundrum we face on wire tapping, alright? So we are in this middle here. So, this is an encrypted conversion and we are curious and we want to know: hey, what can we learn? The truth sounds like a brain dead idea. Why would I do this? This is an encrypted conversation, this is nuts. So, I am nuts and so I have to have some followers who are also nuts are well. So, I need to give them some credit. So, over the years we've had a few people working with me on this started and you will see the transition. And there is going to be a test about that so—later on. Whilst I was at Hopkins, Charles Wright lead this work, this was part of his thesis work and Charles Wright is now at MIT…lab with collaboration Lucas..who is at Google and Scott…who is at (Redjack) which is somewhere in this area. And then some of you might realize that I left Hopkins and moved to UNC and a UNC they had replaced us with three new guys. Much bigger team as you will see as we continue. Andrew White, Austin… White…presentation a few times and if you note here he is at CMU. One of my mistakes but we will come back to that when it's your turn to talk [laughter in the background]. So, let's see…Give you more perspective. Think of it this way we are looking at the voice channel and this is encrypted…confidentially. If I am wear the attacker's hat…you have to give me kiddos for this… You have the attackers here and… on the wire. So what normally happens is that you have a voice signal? And the voice signal usually goes through some (Kodak). And I will talk about what that (Kodak) does and some compression and its going to use some real time transfer protocol…encryption is going to happen here…and its going to send it over the network, right? And you get these and usually about 20 milliseconds of audio is compressed into one frame and sent as a packet along the wire. So we are sitting here and seeing these packets and saying what should we do? Okay? So let's take a look at why this information leakage happens. So, we as designers I think overlook two important design decisions which happens over and over again…this is the standard tradeoff that we pay. One is move towards variable …encoders…compress different sounds with… We do this for efficiency it is very efficient...which is an example of codec cited… in particular uses nine different…which uses…and 21 bits…mode, okay? I forgot about…So we have compression in one end. And then we have encryption. We use what we can…okay? So it's for efficiency. So we have a packet. This is a real time conversation and so we can't buffer stuff. And for encryption, we are just going to encrypt the packet and send it along its way. Okay? So we have this interplayed. And why is this interplay important? It's because we have—This has an unintended interaction. So on your router or () you have a high level or viewer list. You have the—The signal is compressed and then the compressed signal is encrypted but if you notice that if I was looking at the encrypted packets on the wire the packets reflect properties on the input signal. Okay? Let's look at a closer example here. So, here is an audio away form. Here is the output from the codec so bit rate on the y-axis…and here is the output of the wire so if you are looking on the wire. And so, the y-axis is the packaged size and bytes. So, something should jump out at you. The last two plot there is a significant sunken similarity between the last two, right? So we want to go backwards, right? So this is what the work is about, okay? So, okay let's put more context. So, where do you think this combination of variable bit rate encoding and stream cipher is supported? Well, I found out last week ago…actually uses speak the same coding that we use right?… when you use serial… entire utterance is sent back to the server. Alright? So your voice is relayed on some server somewhere. And so in that case, it is send over HDBS connection. Yahoo messenger, Microsoft communicator they all use support BBR and…ciphers and Google Talk are the same and Skype as well. Skype uses its propriety. So, we will get back to Skype later in the presentation and in terms of encryption it also uses a stream cipher which its proprietors who we don't know much about…it is believed to be our variant of our C4, another stream cipher…software PPX like Asteriks…right? So, it's should—The combination should be out there. So, how bad is this leakage? Is the question we get all the time. So, back in 2007 we showed that this is Charles Wright the Hopkins group it is significant enough that we can determine the language of the encrypted conversations so whether this is a phone call in Spanish, French or English, etc. Okay? So in 2008 we should that…spot me if you can we were unable to uncover spoken phrases and encrypted voice sessions. So, this you have to know what you are looking for…get some intuition about a little bit later. There is a lot of work that has happened since then. Several papers that redone our studies and quite a few other twist on this and including some work on…showed on a different assumptions and on some conditions they might be able to tell you about the speaker, okay? So a lot has happened since then. And Keith asked me to come here so something happened between this point and this talk. Alright? And we are going to talk about what happened along the way. Okay? So in respect to back to the original text…that this leakage is not a big deal, alright? So in the first case, the language of the call can be determined in a simple—In a much simpler way by end point analysis so if I see a call between Mexico and Spain—Hey, that is most likely in Spanish. I don’t need to do all the stuff that we did here. And second attack require this…knowledge. You have to know what you are looking for. So, in our case look for the example whether the phrase attack at dawn is there…search basic technique is it likely that the utterance attack at dawn... I have a thought. I see Kevin shaking is shaking head and Bobby… I see this is not realistic. And we—We got a lot of that pushed back. This is great but this is not that realistic. From the pure academics, we also got—This is not that new either which is true everything has been done before. Trust me…And there is this really nice piece of work done by Kevin McCurly…where he noted that. We have to be careful... In the crypto communities…the communication models that we are using don’t really capture the power of the adversaries so he warned about these types of stuff…there is no…warned about this combination is something that you have to pay attention to. From the industry point of view, Phil () was a designer of PGP…project…He was the only one that was trying to encourage others to move away from VBR new, non-VBRB based codecs right he said this might be a problem go back to VBR based codecs. I think that there is a bit of conflict of interest…there… is the …project…pretty much the reaction of the community is that this is not a big deal. So go back to the slide where 2008—Late 2008 there is a late transition from Hopkins to UNC but something more important happened around the time…been around that time. And to sure that my program managers, I was working on this stuff during the transition. Okay? We were thinking about the…problem. So, back around 2009 uh oops—Wait, there you go 2009…Okay, Skype, for the first time released its codec which is a VBR based codec. And the idea was for…specifications to relase