(TS//SI//NF) User's Guide For PRISM Skype Collection

August 2012

POC For Document: [redacted] S3531

1. (TS//SI//NF) Introduction

a. PRISM Skype collection can be identified by the following:

a.i. SIGAD: US-984XN (this SIGAD is used for all PRISM collection, not just Skype collection).

a.ii. Case Notation: P7*, where the * represents a wildcard.

b. Sustained Skype collection began in Feb 2011 against "in and out" modes; this is the mode in which one end of the call is using a Skype application, and the other end is using a landline, or cell phone, without a Skype application. This is audio only. PRISM collection added exploitation against peer-to-peer Skype applications in July 2011. In this mode, both/all users are using the Skype application. This can be a mixture of audio, video, chat, and file transfers. PRISM collection can be from both modes of communication; however, there is no distinction made in the Case Notation.

c. PRISM collection is available for "surveillance" tasking in UTT. There is no "stored comms" or "search" capability for tasking.

2. (TS//SI//NF) Questions on tasking Skype Selectors to PRISM:

a. Can I task something other than SkypeUser?

a.i. No - this is your only option for Skype selectors in PRISM. You cannot task skypeMailToken.

b. Do I need to include a V-Series zipcode?

b.i. Yes. When you submit FAA Skype requirements, please remember to include a V-series zipcode – otherwise, voice content will not be routed to NUCLEON.

Derived From: NSA/CSSM 1-52
Dated: 20070108
Declassify On: 20361001

c. Can a Skype user have more than one skypeMailToken?

c.i. Yes - they can link 2 or more email addresses to their account... each has a corresponding skypeMailToken. You can query DECODEORDAIN to find out what e-mail address they correlate to.

d. Is there any benefit (or harm) to tasking both the SkyUser and SkypeHash for the same user to PRISM?

d.i. Tasking a SkypeUser automatically tasks the SkypeHash to UTT. The only time you need to explicitly task a SkypeHash is when you don't know what the correlated SkypeUser is (i.e. querying DECODEORDAIN with the SkypeHash fails to return a correlated SkypeUser, and after you've contacted [redacted] and they fail to find a correlated SkypeUser).

3. (TS//SI//NF) Questions on locating Skype PRISM data:

a. How do I find Skype data in PINWALE?

a.i. Best way to query in PINWALE is by DecodeOrdain enriching your target's Skype username, as follows:

a.i.1. Enter the Skype username in the PINWALE query form, right click on the username that you've typed-in and click "DecodeOrdain" then "Apply"

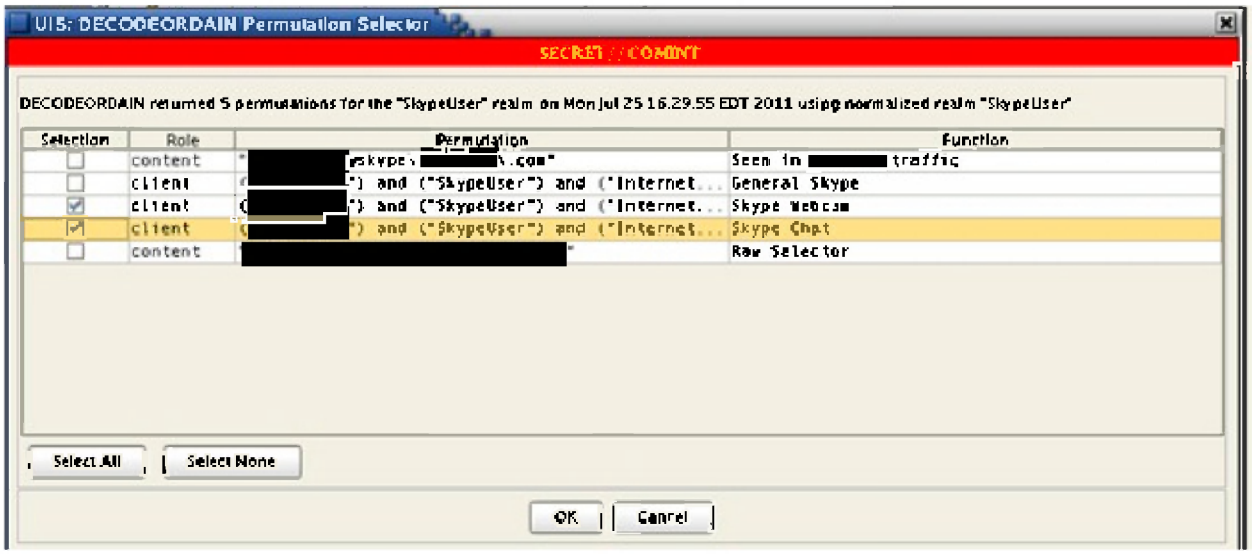
a.i.2. Pick "SkypeUser" from the "choose realm" pull-down list and click OK

a.i.3. You'll now see a screen like the one attached. Pick which ever permutations you want (or just leave them all checked to get everything), and click "OK"

a.i.4. submit your search

b. How do I locate Skype collection using DECODEORDAIN?

b.i. Attached is a screen-shot of how you can use DECODEORDAIN to find Skype video & chat content.



b.i.1. By selecting the "General Skype" permutation, you will query all Skype PRISM traffic (regardless of whether there is any content or not)

b.i.2. By selecting only the "Skype Chat" permutation, you will narrow the query to only Skype PRISM traffic that contains Instant Messaging content

b.i.3. By selecting only the "Skype Webcam" permutation, you will narrow the query to only Skype PRISM traffic that contains at least one visible video stream

b.ii. Please note that when MARINA says that a SkypeUser webcammed with another user... that event *did* actually happen - but the content may not be visible in PINWALE (because part of the video stream may have been lost, prohibiting us from being able to reconstruct the video). If you want to look for PINWALE documents that have webcams that are actually visible, you can use the DECODEORDAIN-enabled PINWALE query option, described above.

c. My Skype PINWALE cut says "PRINTAURA Skype" – what does that mean?

c.i. The display name is “PRINTAURA Skype” for both the UIS Text Presenter and the DNI Presenter. The DNI Presenter and UIS Text Presenter both have the capability to display SKYPE; Chat, File Transfers, and Video. The DNI Presenter has a single combined display for SKYPE webcam that allows the user to play the video and audio together in a synchronized mode or individually. Within UIS, the user will need to use UIS Hotzone to play the corresponding audio cuts.

d. How do I search for Skype data in NUCLEON?

d.i. You can query by case notation (P7* where * = wildcard), or you can query by username in the TELEPHONE_NUMBERS field (note that if your target has a long username, you may have trouble searching by username due to length limitations in NUCLEON... but you won't have any trouble searching by case notation in NUCLEON).

d.ii. Another way to pull up the audio in NUCLEON is to pull on the zipcode that the audio was sent to. If you pull it up this way, the Skype audio displays along with the rest of the FAA audio that was sent to the same zipcode.

e. Where's the other side of my voice cut in NUCLEON?

e.i. CES has added a great new service for OPIs on PRISM (and FBI FISA) voice collection: all voice cuts will now be autopaired and you will not have to hunt down the other side of phone calls. You will automatically be presented with both sides of the conversation in NUCLEON. Previously, analysts would have to search through NUCLEON to find both sides by looking at the targets and the timestamps to guess which ones are paired.

f. Where's the audio to go with my Skype video?

f.i. Within the DNI Presenter (DNIP), the user can utilize the "View Associations" service to find the associated NUCLEON audio of the PINWALE document or find the associated PINWALE document of the NUCLEON audio. In other words, you can find the MAM and/or TAM associations to any associated Skype data, and then display them within the DNIP Skype combined display or the DNIP Composite display. Also from the UIS Text Presenter, you can launch the DNI Presenter to utilize this "View Associations" service.

g. How Do I Tell When Stored/Search Comms Arrive From PRISM?

g.i. There is no collection capability for Stored/Search Comms in Skype PRISM.

h. Why do I receive multiple copies of Skype chat sessions?

h.i. You might get chats in segments and then get the whole chat in a third collect. This is how Skype works. Depending upon what your target is doing, a copy of his chat history can be sent in-bulk (which can span multiple chat sessions). If you target, for example, has 3 separate chat sessions with another individual on his laptop, then logs-into his Skype account on his desktop, the chat-history of those 3 separate chat sessions will be transmitted from this laptop to his desktop so that both his computers have a log of the whole conversation.

i. What does "0 Frames Available" mean in my Skype video results?

i.i. Whenever you see this in PINWALE: "Stream 1 has 0 frames available" or "Stream 2 has 0 frames available", it means that we know that a video-feed was taking place, but we didn't get enough traffic to reconstruct any video frames. Either the target was unable to get a video feed up, or too much traffic was lost during the collection process to reconstruct any images.

j. Why are so many e-mail addresses correlated to one Skype user account?

j.i. These are probably all valid associations. Skype does not require that a user enter a "valid" email address when creating a Skype account (which establishes their skypeMailTokens). A simple email address like "abed@hotmail.com" is likely to be randomly picked by a large population of Skype users when creating their accounts... so it's not unreasonable to find a lot of correlations. You will likely find a similarly high correlation count for something like "abc@hotmail.com" or "asdf@hotmail.com" or "bob@hotmail.com". These are not false hits... these users *did* enter this email address when creating their Skype account... but a lot of them probably did so to maintain their anonymity by typing in something simple & random, rather than their true email address.

k. What's wrong with the timestamps in my Skype data?

k.i. This is due to 2 reasons:

1) Skype sends chunks of old conversations to synchronize conversation histories across multiple Skype installations. DNI Presenter & UIS attempt to detect "historic conversation dumps" and label them as such (but keep in mind these conversation histories can be quite old... days or weeks - so it's important to pay close attention to the timestamps).

2) There are 2 timestamps displayed in UIS & DNI Presenter for Skype. The first is the collection time, and the second is the target's local computer time. DNI Presenter displays them properly as "Collection Time (GMT)" and "User workstation time (GMT)". Please note that NSA has no control over the "User workstation time"... a target could decide to set his local computer's clock months or years off from the true time.

l. Why do I sometimes have to piece together a lot of chat sessions from Skype?

l.i. The reason why you're seeing pieces of Skype chats in different documents is for 3 reasons:

1) the data is too big, and is being fragmented into multiple documents (this isn't that common, however)

2) Skype sends chunks of old conversations to synchronize conversation histories across multiple Skype installations. These old conversations are sent piecemeal across multiple chat sessions.

3) Every time a Skype user disconnects from Skype, and reconnects again - they being a new session. It's common to see Skype users disconnect & reconnect frequently whenever Skype stops performing well for them (i.e. if their video or audio messes up, it's common to see a Skype user disconnect & reconnect to get it working again). Every time they do this, a new PINWALE document is started.

Items 2 and 3 are based entirely on target behavior (it's not a collection problem, and is not being introduced by NSA).

m. Why does my Skype playback get out of sync?

m.i. The main issue is data-loss causing the audio to skip and fall-behind. CES is working on a few techniques to help keep the audio in-sync between channels and with the video (the video should always be in-synch, since we timestamp each video frame).

4. (TS//SI//NF) **Questions on miscellaneous issues:**

a. How do I find out the e-mail address of my Skype target? The collection only gives me the Skypeuser, SkypeNode and SkypeMailToken.

a.i. Email addresses are not included in Skype collection... skypeMailTokens (which are derived from email addresses) are. You can search MARINA for skypeMailToken - to - SkypeUser correlations to detect a username change, in order to put new SkypeUser selectors up on FAA coverage. You can also transform any email address into a skypeMailToken using DECODEORDAIN. DECODEORDAIN can also transform some skypeMailTokens back into email addresses. For cases where DECODEORDAIN cannot transform a skypeMailToken into an email address - please send an email to [redacted] and request that they attempt to recover the email address from the skypeMailToken (however, please only do this if DECODEORDAIN fails to return the email address).

b. Can we collect Skype voicemail?

b.i. If a target connects via a peer-to-peer connection, we will not be able to collect any voicemail he receives or sends. If a target connects via a PSTN connection and uses his In/Out number to communicate with another user's

In/Out number, then we will collect voicemails. Both of these scenarios only apply to surveillance collection. We do not have any stored comms collection from Skype at this time.

b.ii. I/O is a telephone number assigned by Skype. You can use either your Skype ID or your I/O number to make/receive calls. If you're only making P2P calls you don't need an I/O assigned. You only need an I/O number if you want to make calls to/from the PSTN. Note you can also make I/O to I/O calls, which must hit a PSTN switch of some type.

c. What happens when both sides of a Skype call are tasked selectors?

c.i. PRISM should collect two copies of the communication. This is true unless one of the users' Skype communications falls into one of a couple of technical limitations which prevents us from currently collecting via PRISM.

d. Why did I receive small and incomplete files from Skype collection?

d.i. The first 256 bytes of a document for transfer are sent so that the Skype client can figure out what type of file it is (a PDF verses a JPEG). It shows-up on the user's Skype chat window as a message that says something like "so and so is attempting to sent you a file... click here to download... etc.". If the user doesn't choose to download it, then you will not receive the entire file, but only this first nibble of data, which might include the filename.

d.ii. Also, due to known limitations in our collection process, some larger size files being transferred by targets may be not be fully collected. There is no stored comms capability to go back and retrieve these files.

e. How does the Skype Credential CRC work and what will I see in collection?

e.i. The way Skype works, users (in the raw data) are represented by their Skype Credential CRC (SkypeCredCRC). Usually the data contains enough information such that we can figure out what SkypeUser maps to the SkypeCredCRC... but sometimes this isn't possible. SkypeCredCRC's are temporary identifiers... so they're not targetable. Normally analysts don't

have to worry about SkypeCredCRC's, because CES Protocol Exploitation is able to map them to SkypeUser's... but there are a handful of edge cases where they don't have enough information to do so. Also, a SkypeUser can have more than 1 SkypeCredCRC associated with them. So - the possibility that a SkypeCredCRC shows-up in your collection and is associated with one of the already-existing communicants in the collection is a very possible (although it's not universally true). CES only "holds-up" collection for a certain period of time before they give-up trying to figure-out how to map an unknown SkypeCredCRC to a SkypeUser (12 hours would be too long... that would build-in a 12-hour delay into all Skype collection before you got to see it). If you run into a SkypeCredCRC and cannot figure out who it is, just send an email to [redacted] and someone (either in PE, or SSO) can try and figure-out if we (or the FBI) has learned the correlation after-the-fact. It is not possible for you to see the correlation in the raw data. Encryption is used to mask the correlation.

f. How does Skype work?

f.i. Skype is a true "peer to peer" service... communications are not routed between central servers - they are only routed between Skype users (i.e. "peers"). NSA receives Skype collection via PRISM when one of the peers is a FAA/702 tasked target.

There are 3 collection scenarios:

A) target --> other_person

B) other_person --> target

C) target --> same_target

Scenario A and B occur when a target is communicating with another individual. This communication may not be very interesting (i.e. it might not be an IM, Voice, or Video... it may just be a request to retrieve that individual's Avatar to display)... but there was *some* type communication between these two users, and the target is a party to that communication.

Scenario C is an interesting case -- it implies that the target is talking to himself. In fact, he is... This scenario occurs when the target uses 2 different devices (i.e. laptop and mobile phone, or laptop & desktop, etc.), and is logged in from both devices at the same time.

These 2 devices will automatically "synch" with each other, and transmit information between them (including past communications that the target was a party to).

Now - when it comes to multi-parti communications (conference calls, multi-person IMs, etc.), Skype is still a peer-to-peer protocol (1-to-1). Skype creates a mesh-network, where users are connected together through multiple peer-to-peer links. Instant Messages sent to this group of meshed participants can be routed *through* any participant. For example, if A, B, and C are chatting together in a conference chat, the peer-to-peer network may look like this: A <--> B <--> C. In this scenario, IM's from A to B and C will all be routed through B. If B is our target, we will see these IM's... but B is a party to these IMs - because they're multi-party IMs (not private IMs between A and C).

5. (TS//SI//NF) **Other Useful Resources:**

a. Instructional Videos on KapShare Learning Portal

a.i. Go to this link:

[redacted]

and type in “Skype” in the Search box. You will see a list of short videos which describe different aspects of Skype exploitation. As of Nov 2011, the list includes a primer on how Skype works, how to analyze Skype metadata, Skype in XKeyscore and other topics.

b. Email the PRISM technical help team with any Skype question: DL BL_PRISM_HELP.

c. Visit the PRISMFAA web site and search the FAQ section for Skype answers and tips.