TOPIC: Malware

Ivan: We’re going to go over Bret’s proposal that we discussed at the F2F. He’s proposing a new Malware Analysis SDO in addition to the malware SDO. The new Malware analysis SDO allows for versioning as well.

Allan: What’s the definition of Analysis? Is the intention that people can create instances without analysis? And the instance would be an example of the object (specific hash)

Ivan: Agrees. The malware SDO would capture the metadata/labels and sample information. Malware analysis would capture the static/dynamic analysis.

Ivan: we talked about putting high level information on the malware object that could be a summary of analysis information

Allan: Could this be two ways of doing the same thing? (using analysis SDO vs capturing some high level info on the malware instance object)

Bret: I want to get rid of nesting inside the object so other entities (even inside your own org) can add additional analysis going forward. Some pieces of metadata might need to be bubbled up to the instance (rather than the analysis object). There are still things that need to be changed/modified

Allan: I’m not saying that it’s wrong, but best practices are often the best way to solve this. Might need to have best practices.
Trey: STIX is an interchange format not a data model

Ivan: It would likely have some duplication of data, but hopefully that would be things that were consistent from run to run, and different things could show up on each individual run

Jeff: should we be showing the malware analysis directly linked to the observed data (or is that option 1`), since this proposal does lend itself to 1`

Ivan: we’re not ready to have the conversation about observables.

Bret; let’s get high level agreement on malware first

Sean: we agree with creating a separate object for analysis. Malware object is everything we know about the malware, but each analysis is separate. You’ll know what each run discovered separately

Sarah: I would argue that you can have an instance without having the analysis, like if you get a malware report but you didn’t do the analysis yourself.

Allan: I think we’re all roughly agreeing. The comment I made about duplication is different then how it’s being discussed. How would a new person coming to the spec interpret these two possible ways of doing this? Would there be confusion? We just need to clearly define in the spec when you would use the instance object and when you would use the analysis object.

Sean: Agrees.

Ivan: The text is notional, so we would need to make sure that happens. Examples will help.

Bret: The proposal was at the high level, we can get to the details going forward.

Ivan: I’m hearing rough consensus about having a malware analysis SDO, if not all the details. In the way the proposal is currently, you might lose the context of actions. Previously we had a dictionary of key/value pairs to capture the actions. From the F2F we had rough consensus to keep the key/value approach, but we could do something in the future once we have actions.

Sean: The question is broader. How do we represent the relationship between malware and what we know about it. One option was separate relationships to each of those thing. The other was that they could be part of the object (static or dynamic) could be done as a dictionary approach. The approach should capture both features and actions. We could use the dictionary approach for multiple things.

Ivan: I think there are 2-3 possibilities. 1) Keep the current key/value approach (had consensus before. 2) figure out how we’re doing actions and create the corresponding structures (more
extensible, but is more complicated and will take a while). Do we want to do it quickly or do take more time to get it out the door?

Sean: Actions are more important to do right, but it would put us off. We should probably deal with actions later.

Ivan: Actions won’t be quick because of all the nuances. We need to decide approach do we want to take. The F2F consensus was key/value

Bret: I can go work with people to take another pass at the proposal. How malware works with cyber observables needs to be addressed after the TC decides how we are going to deal with observables. If the TC decides 1` is the way to go, then a special type of relationship might work, but in the meantime we could put a flag/dependency

Gary: If we have a specific way to do actions inside malware, which won’t correspond to how we do actions elsewhere, then we’re going to have a large breaking change later. I don’t see that as a good approach. But I still want to see malware out the door out quickly. I don’t think we’ll need to have all the actions defined in order to roll them out. It isn’t all or nothing.

Allan: What are the two options?

Ivan: Key/value pairs and doing actions

Allan: We already have the problem of doing things inconsistently (end result vs action as it happens?) A unified way of doing actions could fundamentally change everything.

Sean: I don’t’ think it fundamentally changes everything. But it lays the groundwork for things to be done consistently.

Trey: Can I ask a question? We’ve been working on malware for two years. What does a minimally viable done state look like?

Allan: From my perspective you want the information so you can detect, block, respond. So as long as the data is conveyed in a machine readable way, then it ‘works’. It doesn’t need to be perfect, but it does need to be consumable. We need to find out what we can’t convey right now, then fix that. If it can be conveyed even if it isn’t perfect we should stick with it.

Trey: Looking at threat feeds, one of the things I’ve noticed is malware/threat reports with no differentiation between artifacts malware family foo and malware instance foobar. Maybe we should aim for minimally viable.
Bret; I think we’re trying to find something usable for those that create malware data. If threat feeds don’t have it, it’s because the standard isn’t there. We get one more chance with malware before people do their own thing.

Jeff: I don’t think the action object is that crazy. I made a simple proposal. As long as we can capture a subgraph, that’s fine. If we don’t do actions, but we do 1 we could use relationships without much loss.

Ivan: I don’t want to get into actions now. But if we want to have summary data at the instance level, and we need a separate actions, we could reuse our key/value pairs in the summary data

Sean: What’s minimal is the ability to describe malware analysis that occurred and then the static/dynamic features. What do we know about it? What do we know about that instance? What do we know about that family? How is that tied to other SDOs?

Trey: Right now we have a stub. We could snap our fingers make it 3x better and still not achieve what you want.

Sean: But if we don’t get to what we need, I’d rather leave it alone.

Trey: So leaving it a stub is better?

Sean: Yes. Because it would just be a “better” stub. You can’t convey what we need to describe. If you can’t describe the basic stuff, why are you doing it? Dictionaries isn’t the perfect solution it’s the minimal solution.

Trey; I would like to move that we ship 2.1 with the malware stub as-is.

Sean: Are you being sarcastic?

Trey: Yes.

Allan: We’re all producing malware artifacts. We all spent lots of time to convert from one proprietary format to another. It’s a drain on the community. We have an obligation to come up with a definition that we can use that can help all our customers and the industry as a whole. We need to agree on a minimum set. Should we stick with key/value pairs? Is it good enough? Yes. Is it perfect? No.

Sean: I think for the features, key/value pairs would work.

Allan: Let’s take that as a win and move forward.

Ivan: works for me
Bret: That’s a good proposal, but depends on the rest of the observable outcome.

Allan: I’m just trying to get some consensus and move forward.

Sean: I’d agree to dictionaries. But let Jeff comment.

Jeff: I’d say key/values is a 60% solution, we could get actions, but it’s more useful to get some solution and do actions in the future.

Bret: I think dictionary based approach is a viable outcome, pending observables debate.

Sean: The dictionary is still viable with 1`, just not optimal. I’d agree that this decision does depend on the observables outcome.

Bret: The reason I ask the question is that with the observed data object you can’t link the stuff inside the object. If we can’t do that, we can’t. We could enable that with 1`, but we’d have to revisit this.

Sean: I agree. The other question I have is is that the malware object can represent a family or an instance.

Bret: No, “is_family” is back. I had put it into type, but after feedback changed it back to a property.

Sean: Sounds like we agreed. Malware Analysis is separate SDO. Dictionary should work. “Is_family” is back. Those top pieces are great.

Trey: Where are we? What are next actions?

Ivan: We need to flesh out the proposal, maybe Bret and I can do that. We need to think about things/revisit some things. Then regroup in a week or two. We need to have the observed data conversation soon.

Trey: One of the major blockers was to discussing observables was patterning. Also how malware would use the observables. Can we do the modeling now for malware for 1` and 7?

Sean: Not sure about the malware models, but I’ve gone through parts 1-5 to show what they would look like with 1`. Still trying to finish it up, but I’ll be able to share soon. I think we’re close to being able to have the “more data”.

Allan: I’m interested in the changes to the spec and also to the interop docs. What I think needs to be clear is what is the business justification to change from the current way? Make it clear
what can’t be done currently and why 1` is necessary. I need to hear the high level justification. I think we need to preface the changes.

Sean: I’ll have a preface with the changes and why. We were already asked to make the spec changes, so I spent the time to do so.

Allan: If the justification is strong, then it’s easier to move into what the spec changes would be. We need to look at what the implications would be. I don’t want to look at spec changes if I don’t understand why I need to do it.

Bret: We need the evidence based approach. What’s in the 80/20 that we CAN’T do. It’s not about what’s more elegant. That ship has sailed.

Allan: If we had time in advance of the meeting to consume the changes it would help drive to agreement. We might get a better outcome.

End Meeting