5-2020

The future of work now: Cyber threat attribution at FireEye

Thomas H. DAVENPORT
Babson College

Steven M. MILLER
Singapore Management University, stevenmiller@smu.edu.sg

Citation

DAVENPORT, Thomas H. and MILLER, Steven M.. The future of work now: Cyber threat attribution at FireEye. (2020). *Forbes*. 1-5.
Available at: https://ink.library.smu.edu.sg/sis_research/5134

Publication: Forbes Online
Date: 28 May 2020
Headline: The Future Of Work Now: Cyber Threat Attribution At FireEye

The Future Of Work Now: Cyber Threat Attribution At FireEye

Tom DavenportContributor
Enterprise & Cloud

Thomas H. Davenport and Steven Miller

One of the most frequently-used phrases at business events these days is "the future of work." It's increasingly clear that artificial intelligence and other new technologies will bring substantial changes in work tasks and business processes. But while these changes are predicted for the future, they're already present in many organizations for many different jobs. The job and incumbent described below is an example of this phenomenon. It's a clear example of an existing job that's been transformed by AI and related tools.

Steven Stone is Director of Adversary Pursuit at FireEye, the intelligence-led security company. His group is part of the company's Advanced Practices team that focuses on determining the identity, actions, and next steps for cyber threat groups actively operating against FireEye clients. When a new cyber threat group or cluster shows up on the FireEye global cyber threat tracking "radar screen", it is identified as an uncategorized group or cluster (UNC). If Stone's team is able to determine that this UNC is the same as a previous threat cluster they have been tracking, they can merge the two groups and draw upon all their existing knowledge to anticipate what will happen next. They can also see how that cluster is evolving.

However, as Stone comments, if they do not know whether the perpetrating cyber threat entity is a familiar entity, "Then it is as if you are feeling around in the dark. You do not know where to look. It is hard to focus your response efforts. It is more difficult to anticipate what the perpetrator will do next. "

He elaborates, "Until 2018, our method for comparing UNCs was purely manual, and corresponding to that, our approach to the decision of whether or not two UNCs could be merged and considered to be the same entity was purely manual—and required the focus of our top experts."

An Intelligent Tool for Similarity Somparisons of Uncategorized Cyber Threat Groups

Because FireEye tracks thousands of UNCs and sizable numbers of other threat groups, it is impossible for even a team of expert analysts to keep all them in mind at once, and even more difficult to make these comparisons over long periods of time. Said Stone:

"This is the problem we threw machine learning (ML) against. We wanted intelligent, automated tooling to help us systematically and objectively make this comparison of how similar one UNC is to all other UNCs, as well as to the entities in other attribution categories." FireEye network, endpoint, and email security controls deployed across the globe are built to allow massive amounts of telemetry to flow back to a central source, where it can be centralized, standardized, automated and scaled. Stone notes that this approach has been the key to the company's success, as it can use the telemetry data across global client sites to monitor the cyber threat situation across the entire world.

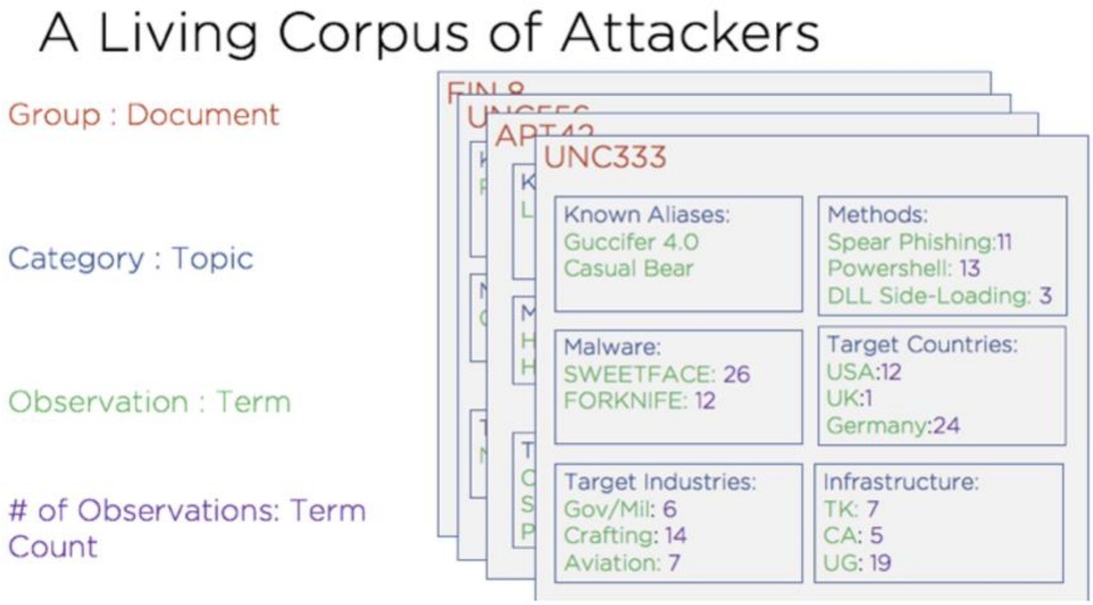
However, having vast quantities of telemetry information and knowing how to systematically harness it for complex comparisons of threat clusters are two different matters. Stone explains:

Publication: Forbes Online
Date: 28 May 2020
Headline: The Future Of Work Now: Cyber Threat Attribution At FireEye

“It is the number one challenge my team deals with. How do you do the type of highly detailed, complex work we do in a ‘bucket’ of data that is this big?”
Domain Experts and Data Scientists Team Up to Create the New ML Tool

Stone and his team came at this dual challenge of comparing UNCs and harnessing the vast amount of FireEye global telemetry data required for these comparisons from two ends. At one end was his team of highly specialized cyber threat attribution analyst with deep expertise in identifying, tracking and pursuing UNCs. On the other, there was the FireEye Data Science team. While the latter did not have the same level of in-depth domain expertise in cybersecurity threat analysis, they were able to build, test, and validate the ML models against available datasets—none of which Advanced Practices had as a capability. Based on their domain expertise, the cyber threat attribution analysts identified almost 50 important dimensions of a cyber threat. They worked with the Data Science team to update and reorganize the data set on all the UNCs they had been tracking in order to describe each UNC in terms of these dimensions.

Even with the data on each UNC organized in this way, there were still too many dimensions and associated metrics. Continuing their collaboration with the FireEye Data Science Team, the analysts developed a modeling framework based on established methods used by the Natural Language Processing and ML research communities to assess the degrees of similarity across different text documents.

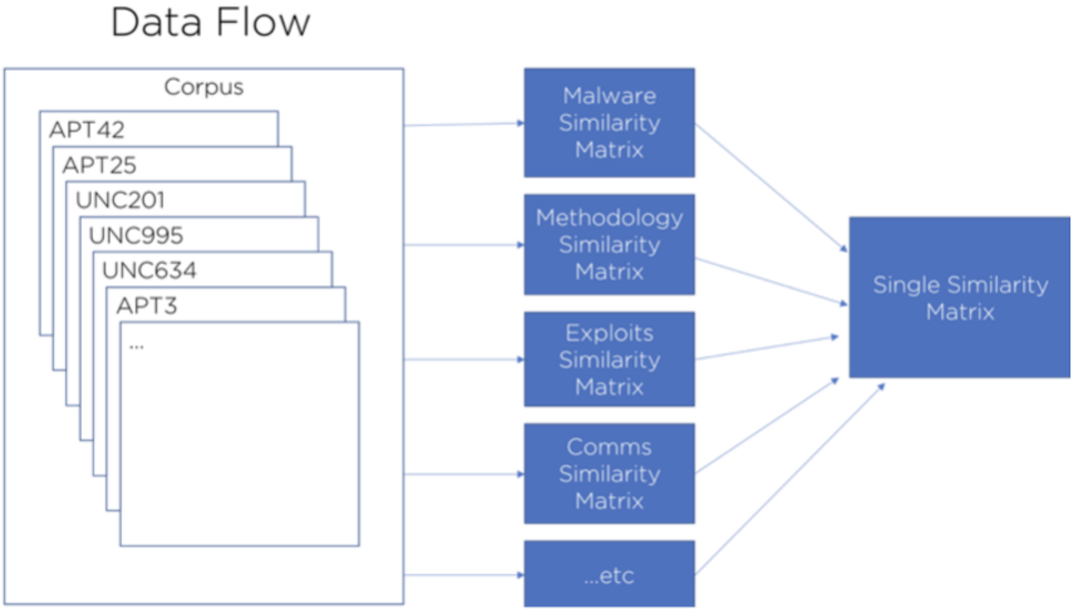


Their big insight was the analogy of mapping their specific need to assess the similarity of cyber attack threat clusters to ML-based natural language processing (NLP) methods for automatically assessing the similarity of text documents. This insight would never have occurred without the intensive back-and-forth interaction between the threat analysis domain experts and data scientists.

Publication: Forbes Online
Date: 28 May 2020
Headline: The Future Of Work Now: Cyber Threat Attribution At FireEye

ATOMICITY Supports Both ML and Human Learning

ATOMICITY was the tool the team developed for evaluating the similarity of cyber threat clusters. To evaluate and validate it, the team used it on historical information to look at all the previous decisions FireEye expert threat attribution analysts had made for merging unidentified threat clusters. These prior decisions were based on human expert determination of the merging.



As Stone’s team continues to use ATOMICITY, the ML system and human experts help train one another in ways that improve learning on both sides. The new ML-based ATOMICITY tool has changed how the team does threat cluster similarity comparisons, and more broadly changed how FireEye works across the company.

Because ATOMICITY can automatically run similarity analysis across thousands of UNCs, FireEye can analyze much more frequently. This has given the company new visibility into how their entire “universe” of UNCs are moving towards or away from one another over time. This new view provides powerful insights to Stone’s team and other FireEye analysis teams to understand the evolution of the global cyber threat landscape.

Stone emphasizes that ATOMICITY does not replace any expert threat attribution analysts, but instead augments and expands what they do. He clarifies, “We do not allow the ML-based tool to make the critical decision of whether or not two [UNCs] should be merged. Only our team of expert threat attribution analysts can make that type of decision, using the supporting analysis information from the ATOMICITY tool.”

ATOMICITY has enabled FireEye to automate selective, essential parts of their work processes in ways that enable the expert analysts in the Advanced Practices group to free up substantial time and mental capacity for doing special investigative projects where humans are far superior to automated ML models. Stone notes:

“The use of this ML-based tool in our group has paid for itself many times over.”
Expanding Use Cases for ATOMICITY

Publication: Forbes Online
Date: 28 May 2020
Headline: The Future Of Work Now: Cyber Threat Attribution At FireEye

As FireEye gains experience and confidence in using ATOMICITY, the company has found new use cases for itself and its customers. For example, cyber threat clusters sometimes drop “off the radar” and then suddenly reappear. It is also common for a new cyber threat entity to unexpectedly appear “on the radar”. In both situations, a wider set of FireEye analysts beyond Stone’s group are making use of outputs enabled by ATOMICITY to probe for possible explanations. In addition, FireEye now includes content in special communications with their customers that incorporates analysis on UNCs from the work of their threat attribution experts using ATOMICITY.

Prior to ATOMICITY, FireEye was reluctant to discuss this type of information with customers, as it was challenging to quantify, justify, and explain the approaches underlying the analysis. However, using ATOMICITY changed their thinking. Additionally, the purely manual process left little bandwidth for customer communications. FireEye now has the confidence and capacity to share some of their assessments on UNCs as they have a much stronger foundation and methodology for their assessments.

Can ML-based models be used more directly to support the analysts in predicting what the threat entity will do next, or in a future time period? Stone sees this type of prediction as beyond the current state-of-the-art, and to the best of his knowledge, no commercial or government entity is known to have such predictive ability. FireEye wants to eventually work out practical and explainable methods for predicting what a threat entity will do in the future, even if it is a behavior that has not yet been observed in the existing data on that entity.

Stone believes that ML-based tools are most useful in helping his team with high-volume work. He elaborates:

“Purposely, our strategy has not been to use ML-based tools to search for ‘the needle in the haystack’. …Our experience is that human experts are far superior—now and for the foreseeable future—for doing this type of early stage, ill-defined exploration, and that the data driven ML-based tools are more productively deployed across our broader based areas where we are deluged with continuous streams of data.”
Another important aspect of ML development and deployment strategy for FireEye is “no auto-magic”. The team will not use ML-based systems to generate analysis or make recommendations where they cannot understand what the system is doing or how it arrives at the conclusions of its analysis.

Reflections on Big Changes

We asked Stone to reflect on the big changes he has experienced during the course of his professional career, starting out as a military intelligence analyst, and later as a cyber threat intelligence analyst. He shared the following story:

“Years ago, when I would sometimes teach courses at one of the U.S. military schools for training intelligence analysts, we would teach them how to investigate many different sources of information to find one useful piece of information. The assumption was that nothing was available, and by cleverly finding one piece of useful information, it was an intelligence breakthrough… I still sometimes teach at that same school for military intelligence analysts, and the entire game of intelligence analysis and similarly cyber threat analysis has changed. It has gone from how to find any useful pieces of information—on the assumption that so little information was available—to how to sift through vast amounts of information and know what to discard in order to discover and retain the subset of it that really matters for the task at hand. This is an entirely different skillset and mindset. It could not be more different.”

Publication: Forbes Online
Date: 28 May 2020
Headline: The Future Of Work Now: Cyber Threat Attribution At FireEye

Steven Miller is a Professor of Information Systems and Vice Provost for Research at Singapore Management University.