18201 Von Karman, Suite 700
Irvine, CA 92612
949.381.7255
www.cylance.com

Cylance's Response to the Request for Information on Current and Future States of Cybersecurity in the Digital Economy

Cylance appreciates the opportunity to provide information to the Commission on Enhancing Cybersecurity ("Commission") in response to NIST's request for information on *Current and Future States of Cybersecurity in the Digital Economy.* As described below, Cylance recommends that the Commission:

- Consider the use of artificial intelligence (AI), algorithmic science, and machine learning to proactively find and prevent the vast majority (99%) of cyber-attacks before they execute;
- Encourage and empower Federal, State, and Local government agencies with the ability to rapidly pilot, procure, and deploy cutting edge tools that leverage machine learning and algorithmic science; and
- Encourage the widespread adoption of a math-driven machine learning approach to cybersecurity as a best practice.

Application of Artificial Intelligence, Algorithmic Science and Machine Learning to Cybersecurity

Machine learning is a branch of AI that concerns the construction and study of systems that can learn from data. It focuses on prediction, based on properties learned from earlier data. Machine learning goes hand-in-hand with data-mining, which focuses on the discovery of previously unknown properties of data, so that those properties can be used in future machine learning decisions. By using machine learning with data mining it is possible to quickly differentiate malicious files from safe or legitimate ones.

Machine learning in this context leverages a four phase process: collection, extraction, learning, and classification. File analysis involves the collection of a statistically significant amount of data. File attributes are then extracted from the data to identify the broadest possible set of characteristics of a file. The result of attribute identification and extraction process is the creation of a file genome similar to that used by biologists to create a human genome. This genome is then used as the basis for the creation of mathematical models to determine expected characteristics of files.

Once the attributes are collected, the output is normalized and converted to numerical values that can be used in statistical models that accurately predict where a file is valid or malicious. This is the learning phase. For each and every file, thousands of attributes are analyzed to differentiate between legitimate files and malware. Once the statistical

18201 Von Karman, Suite 700
Irvine, CA 92612
949.381.7255
www.cylance.com

models are built, an application will use them to rapidly classify files as valid, malicious or unknown with unprecedented accuracy and before malicious files have an opportunity to execute. 1.

Using this approach, every file that is analyzed is evaluated using the classification algorithms. More traditional approaches to security evaluate a specific file against a finite list of signatures designed to detect malware based on human analysis. Even if they use some automated techniques, they are limited to creating signatures based on specific parts of files that were previously identified as known malware. Not only is there little to no proactivity possible with these techniques, they simply classify objects that do not match any particular signature as good.

The ability to utilize machine learning-based technologies that analyze files and determine whether they are good or bad based is purely on the attributes of the files themselves and to do that at a sustained and massive scale in milliseconds has the potential to shift the balance between offense and defense in cyberspace.

In order to quickly leverage the promise of this innovative approach to cybersecurity, Cylance recommends that the Commission encourage and empower Federal, State, and Local government agencies with the ability to rapidly pilot, procure, and deploy cutting edge tools that leverage machine learning and algorithmic science. Cylance further requests that the Commission consider recommending that the use of a math-driven machine learning approach to cybersecurity be adopted as a best practice.

1 -The Committee on Oversight and Government Reform's report on the OPM Data Breach recently highlighted the efficacy of Cylance's technology at detecting key malicious code and other threats to OPM. See Committee on Oversight and Government Reform Majority Staff Report *The OPM Data Breach: How the Government Jeopardized Our National Security for More than a Generation* dated September 7, 2016 at page viii.