2018 TAG CYBER SECURITY ANNUAL

VOLUME 2

INTERVIEWS WITH CYBER LUMINARIES

Expert Advisory Research

Dr. Edward G. Amoroso
Chief Executive Officer, TAG Cyber

September 2018

About the Author: Dr. Edward G. Amoroso is CEO of TAG Cyber LLC. Ed recently retired from AT&T after thirty-one years of service, beginning in Unix security R&D at Bell Labs and culminating as Sr. VP and CSO of AT&T from 2004 to 2016.

TAG Cyber Controls - *Security Analytics*

NIKSUN

> "Dr. Pruthi and his team at NIKSUN work hard to achieve network forensics solutions with lossless packet capture," said Dr. Amoroso. "Building capture solutions that scale to over 100 Gbps without losing packets and developing an analytics platform that scales to such data rates are great accomplishments in technology and have been critical to NIKSUN's success."

Brief Overview

NIKSUN develops real-time and forensics-based cyber security and network performance management solutions.

> "Modern enterprise and critical infrastructure protection demand advanced data capture and processing tools [for providing rapid cyber analytics for data in motion]. Few companies understand the requirements for high-speed packet capture and analytics-based processing to detect cyber indicators like NIKSUN, who can support [this] at extremely high network capacity rates."

Key Executives

Dr. Parag Pruthi serves as Founder and CEO of NIKSUN.

> "Dr. Parag Pruthi is not only a great contributor and representative of our industry, but also an inspiration for his vision of how cyber protections must evolve. The NIKSUN story is a great one."

Products and Services

NIKSUN provides a range of network security and monitoring solutions that can be grouped as follows:

- **Cyber Security** – Includes the NetDetector family of packet capture and metadata analysis products. NetDetector is a full-featured network security appliance. NetDetectorLive integrates packet capture, metadata generation, real-time indexing up to Layer 7, IDS (signature and anomaly), and malware analytics. Virtual NetDetector supports cloud deployments. IntelliDefend is a lightweight (notebook size), full packet capture, and forensics/analytics device for branch offices. NetOmni provides a single, unified view of all traffic across the entire network. Add-On Solution Modules are available.
- **Network Performance** – NetVCR is an appliance for network performance monitoring. It includes flow aggregation, analytic support, and other features. Virtual NetVCR supports cloud deployments. IntelliNetVCR is a lightweight, notebook-sized device for branch offices, department levels, and other applications.
- **Mobility** – NetMobility offers real-time analysis capability for EPC, IMS and CDMA monitoring and analysis on a mobile network.
- **Financial** – NetTradeWatch provides end-to-end visibility into financial network environments.

History

Founded in 1997, NIKSUN provides network forensics solutions that scale to over 100 Gbps with lossless packet capture capabilities. NIKSUN, headquartered in Princeton, New Jersey, has sales offices and distributors throughout the US, Europe, the Mid East and Asia-Pacific.

Headquarters

NIKSUN, Inc. Corporate Headquarters
457 North Harrison St. Princeton, New Jersey 08540. Tel: 609 936 9999. www.niksun.com

Providing Rapid Cyber Analytics for Data in Motion

Modern enterprise and critical infrastructure protection demand advanced data capture and processing tools

Dr. Parag Pruthi, CEO of NIKSUN

Few companies understand the requirements for high-speed packet capture and analytics-based processing to detect cyber indicators like NIKSUN. The company has been at the forefront in this area for many, many years, and its founder and CEO, Dr. Parag Pruthi, has been improving technology platform solutions in this area for decades. We sat down with Parag recently to ask him to share his thoughts on how platform design is evolving, as well as how the underlying behavioral analytic algorithms are improving to the point where they can dependably identify real cyber attack indicators in enough amounts of collected packet data.

EA: What types of packet capture features are your customers requesting?
PP: From the time I started working in cyber security and founded NIKSUN, our message has been consistent, loud, and clear – Murphy's Law paraphrased as "the packets you did not have were the packets you needed" is well and alive. The business implications of this empirical rule are profound. That is, the very tools you don't have are just the tools you needed to do the job right. As a result, the number one feature our customers request is this: "Don't lose any packets because I don't know when I will need the one that were lost." However, for NIKSUN, this is a given – zero packet loss at 1Gbps, 100Gbps, 1000Gbps, or whatever rate is desired. The second feature concerns help with the question: "How do I find the needle in the haystack?" To satisfy this request, at NIKSUN, we first index everything, from the packets and the data in those packets to the sessions, and the data in those sessions, to the applications and the data in those applications. Next, we provide a single portal for analyzing all this data from anywhere and at any time. That is, whether the data is in the cloud, in different data centers, in a virtual environment, or scattered across an enterprise over multiple asymmetric routes, it all just needs to be accessible to an analyst in the same way from one place, irrespective of the analyst's physical location or device. Last, a third and an increasingly important feature requested by our customers is

support for easy and fast visualization. That is, they are looking for help with the question: "How do I know what to look for?" Now, satisfying these, and many other requested features, and making it all work while the application landscape underneath you is constantly changing is no easy feat. Having done this well is the only reason why NIKSUN is the solution-of-choice for not only the U.S. Department of Defense, but also any enterprise for which actual results matter more than marketing hype or personal connections.

EA: Are behavioral analytic algorithms efficient enough to perform sufficient processing for real-time networks?
PP: Well, it depends! Some analytic algorithms are efficient and others are not. The distinction lies in understanding their practical use in cyber security. For example, we can perform principal component analysis in real time and build a language to form expressions of those components. Under certain conditions, this method works rather well and can capture known anomalies where signature detection would fail. Also, deep learning methods, such as deep or recurrent neural networks, are in vogue today, and I am often asked if they can be used to find zero days and stop all cyber attacks. On the one hand, machine learning (ML) works well in some domains where classification is rather straightforward and ample training data is available to converge the algorithms at the minima and not get stuck at false valleys. At the same time, without significant commonality in the various attack vectors and the lack of sufficient training data, all behavior analytic algorithms (ML-based or not) need to narrow the analysis using a variety of depth or breadth first search algorithms. The resultant state space can become so complex that it is not possible to do so within practical budgets. Thus, while there exist algorithms that are specifically devised to detect certain anomalous conditions and are amenable to real-time analysis, many of them are unfortunately not yet computable in real-time. At NIKSUN, we develop both real-time and non-real-time expert systems which encompass various algorithmic analytic techniques.

EA: What is the accuracy of typical analytic algorithms in detecting threats on high-speed networks? Is the false positive rate low?
PP: Despite the numerous success stories of purposefully-designed ML-powered artificial intelligence systems – for example, all of Facebook's translations are now completely powered by an unsupervised deep learning system – their applications to cyber security have been less than stellar. One of the main reasons is that in cyber security, a key challenge is the detection of "unknowns" (i.e., threats never seen before) in close-to real-time despite often very weak signals. For various reasons, this is a task at which unsupervised ML does not excel. Outward signs of this mismatch between what the cyber security domain demands and what unsupervised ML techniques are good at are unacceptably high false positive rates that severely limit the use of ML-based AI systems in practice and unreasonably large mean dwell times that all but guarantee that the attackers have the luxury to take their time to achieve their various objectives. The basic problem with using such systems is that once the analysts lose faith in them due to the high number of false positives, they tend to ignore all the generated alerts and fall back on what works for them – performing everything manually. They react similarly when faced with detection times for breaches that are measured in days and months. Thus, the net effect of using such systems can be self-defeating when applied to the domain of cyber security without proper

restraints. However, for carefully designed systems that are applied with the proper restraints, the cyber security domain provides enormous opportunities. For example, when using the fundamental approach (i.e., "the NIKSUN way") of collecting and indexing all the data and combining it with both algorithmic techniques and computer-assisted but human-navigated analysis, the results can be remarkable. By experiencing efficiency gains far exceeding 500% over traditional methods, many of our clients can do significantly more work with fewer people; or in other words, their analysts can focus exclusively on getting the upper hand over the bad actors.

EA: Do you see changes in the mix of hardware and software required to provide advanced analytics at line speed?
PP: Well, it depends on the line speed. At low speeds, software-only solutions will suffice. But at very high rates, software alone on general purpose hardware is inefficient and impractical. Somewhere in between, a mix can be leveraged. My point is that to be able to defend oneself against attacks such as the recent zero days like WannaCry and Petya as well as a host of other more complex cyber-attacks, one must consider all possible known or unknown attack vectors. As a result, one has no choice but to deal with the problem head-on. Basically, advanced analytics at line speed poses three big challenges. We already talked about the technical problem of performing lossless packet capture and simultaneously generating associated meta-data at high speed (e.g., 1-100 Gbps and beyond). Next, today's cyber security is all about close-to-real-time detection and mitigation of nefarious activities, with the added feature of being able to perform retrospective network forensics when needed or required. This overarching desire for real-time solutions upends traditional analytics and requires the collected data to be treated as streaming data where any analytics is based on a one-time exposure to the data (i.e., at the time of data capture). Essentially, batch processing techniques need to be reinvented to work in real-time. A further challenge is posed by the distributed nature of a typical modern enterprise network. In fact, the ability to collect high-velocity and high-volume streaming data at different locations in such environments mandates a fundamental shift in data analytics. The traditional view that "the data has to be moved to where the analytics/processing is done" is replaced by the new insight that it is the analytics that must be distributed (i.e., the analytics/processing has to be brought to where the data resides). Even though these challenges have been known for the last 20 or so years, some people still wrongly think that a simplistic mix of solutions can address them. For example, one popular solution is to have a device classify all the data, such as a firewall, and another one to simply collect packets. The problem with this approach is that doing the real-time analytics or post-event analysis without the appropriate metadata is by and large useless. By the time one is done fetching the packets for specific flows and then reassembling them for analysis, many other events will have queued up and there is no digging out of this hole. At NIKSUN, we have studied this problem very carefully and ended up designing a solution that we optimized in both space and energy, in hardware as well as in software. NIKSUN's Supreme Eagle architecture, with its built-in support for cluster and grid computing, provides exactly the type of system-level support that this paradigm shift in advanced analytics requires. As an all-in-one platform, it offers the basic functionalities for real-time analysis of the type of "hyper data" that it collects. In fact, it is ideally suited for implementing distributed streaming data algorithms that are at the core of any advanced

analytics in support of real-time cyber security solutions. We have advanced this mix of hardware/software analysis so far that we are now exploiting the full power of this type of advanced analytics to harness unprecedented opportunities for both real-time cyber security solutions as well as "back-in-time" analysis. Moreover, by supporting this type of advanced analytics on our suite of virtual solutions, we can offer customers "network monitoring as a service" and enable them to reap the benefits of network function virtualization (NFV) by letting them decide where to perform ultra-high performance packet capture and analytics, when, and for how long. NIKSUN's virtualized software takes full advantage of dedicated hardware and provides scaling in multiple dimensions.

EA: How important is domain knowledge to detect network attacks for applications such as industrial control or IoT?
PP: If the past is any indication of what the future in cyber security has in store, we would be foolish to envision that we will be able to completely replace domain experts or eliminate humans from the loop by leaving it all up to AI to do the job for us. Whether we are concerned with protecting the various systems that control the myriad of different industrial organizations and critical infrastructure networks we rely on in our daily lives or worry about nefarious activities that potentially involve millions of vulnerable IoT devices and can presumably cause havoc at local or global scales, domain knowledge will remain the go-to solution so long as the software for the control is written by humans. Just as domain knowledge is paramount for finding bugs in this software, recognizing how they can become vulnerabilities when used for nefarious activities, and ultimately exploiting them for specific attacks, it is also essential for reverse-engineering a given (unknown) bug from an observed attack. While AI in its current form is ill-suited for both these tasks, domain experts excel in them. At the same time, once the basic mechanisms underlying such "unknowns" have been elicited and are understood and known, the job of detecting future occurrences of the same type of attack in real-time can be left to AI after the successful implementation of suitable real-time analytic algorithms that mimic the steps used by the domain expert to get to know these unknowns. It is in this sense that existing and emerging AI approaches can be fully expected to play a critical role in securing our future networks against cyber attacks. By automating all the tasks that are amenable to automation, we reap the benefits of AI systems by putting ML techniques to work on problems where they reign supreme – detecting "known bad" activities with high confidence and preventing "known good" activities from triggering false alarms. At the same time, this use of AI also frees up the domain experts to focus on work where they excel at – getting to know the unknowns in a gradually diminishing portion of suspicious traffic. It is in this sense that I believe that the holy grail of cyber security – that is, the real-time detection and mitigation of nefarious activities – will for the foreseeable future require human involvement in the form of cyber security experts and their invaluable domain knowledge.