

**UNITED STATES PATENT AND TRADEMARK OFFICE**

---

**BEFORE THE PATENT TRIAL AND APPEAL BOARD**

---

CROWDSTRIKE, INC.,  
Petitioner,

v.

SKYSONG INNOVATIONS, LLC,  
Patent Owner.

---

PTAB Case No. IPR2025-01170  
U. S. Patent No. 11,892,897

---

**DECLARATION OF NENAD MEDVIDOVIC, PH. D., IN SUPPORT OF  
PATENT OWNER'S PRELIMINARY RESPONSE**

I, Dr. Nenad Medvidovic, under the penalty of perjury under the laws of the United States, declare that the following is true and correct based on the best of my ability.

Date: 9 November 2025

Signed:

  
\_\_\_\_\_  
Nenad Medvidovic, Ph.D.

## TABLE OF CONTENTS

|      |                                             |    |
|------|---------------------------------------------|----|
| I.   | QUALIFICATIONS .....                        | 6  |
| II.  | BASES OF OPINIONS .....                     | 17 |
| III. | A PERSON OF ORDINARY SKILL IN THE ART ..... | 18 |
| IV.  | LEGAL FRAMEWORK .....                       | 19 |
| A.   | Claim Construction.....                     | 19 |
| B.   | Validity .....                              | 20 |
| 1.   | Anticipation.....                           | 21 |
| 2.   | Obviousness .....                           | 22 |
| V.   | THE '897 PATENT (EX1001).....               | 24 |
| A.   | Technical Background.....                   | 30 |
| B.   | Challenged Claims.....                      | 33 |
| 1.   | Claim 1 .....                               | 34 |
| 2.   | Claim 2 .....                               | 34 |
| 3.   | Claim 3 .....                               | 34 |
| 4.   | Claim 4.....                                | 35 |
| 5.   | Claim 5.....                                | 35 |
| 6.   | Claim 6.....                                | 35 |
| 7.   | Claim 7.....                                | 35 |
| 8.   | Claim 8.....                                | 35 |
| 9.   | Claim 9.....                                | 36 |
| 10.  | Claim 10.....                               | 36 |

|                                  |    |
|----------------------------------|----|
| 11. Claim 11 .....               | 36 |
| 12. Claim 12 .....               | 36 |
| 13. Claim 13 .....               | 36 |
| 14. Claim 14 .....               | 37 |
| 15. Claim 15 .....               | 37 |
| 16. Claim 16 .....               | 37 |
| 17. Claim 17 .....               | 37 |
| 18. Claim 18 .....               | 37 |
| 19. Claim 19 .....               | 38 |
| 20. Claim 20 .....               | 38 |
| C. Priority Dates .....          | 38 |
| VI. THE ASSERTED PRIOR ART ..... | 39 |
| A. Sabottke (EX1004) .....       | 39 |
| B. Smyth (EX1005) .....          | 42 |
| C. Phillips (EX1006) .....       | 44 |
| D. Nunes (EX1007) .....          | 45 |
| E. Kuperman (EX1008) .....       | 45 |
| F. Marin (EX1009) .....          | 46 |
| G. Bozorgi (EX1010) .....        | 47 |
| VII. CLAIM CONSTRUCTIONS .....   | 47 |
| A. “social connections” .....    | 47 |

|       |                                                            |    |
|-------|------------------------------------------------------------|----|
| B.    | “disclosure date” .....                                    | 49 |
| C.    | “exploitation date” .....                                  | 49 |
| D.    | “likelihood of exploitation” .....                         | 49 |
| VIII. | VALIDITY ANALYSIS .....                                    | 51 |
| A.    | Ground 1: Obviousness in Light of Sabottke .....           | 51 |
| 1.    | Claim 1 .....                                              | 52 |
| 2.    | Claim 3 .....                                              | 65 |
| 3.    | Claim 5 .....                                              | 65 |
| 4.    | Claim 9 .....                                              | 65 |
| 5.    | Claim 10 .....                                             | 66 |
| 6.    | Claim 14 .....                                             | 66 |
| 7.    | Claim 15 .....                                             | 68 |
| 8.    | Claim 18 .....                                             | 68 |
| 9.    | Claim 20 .....                                             | 69 |
| B.    | Ground 2: Obviousness in Light of Smyth and Phillips ..... | 69 |
| 1.    | Claim 1 .....                                              | 70 |
| 2.    | Claim 2 .....                                              | 76 |
| 3.    | Claim 3 .....                                              | 77 |
| 4.    | Claim 5 .....                                              | 77 |
| 5.    | Claim 10 .....                                             | 77 |
| 6.    | Claim 11 .....                                             | 78 |

|     |                                                                      |    |
|-----|----------------------------------------------------------------------|----|
| 7.  | Claim 13.....                                                        | 78 |
| 8.  | Claim 14.....                                                        | 79 |
| 9.  | Claim 15.....                                                        | 80 |
| 10. | Claim 17.....                                                        | 81 |
| 11. | Claim 18.....                                                        | 81 |
| 12. | Claim 19.....                                                        | 82 |
| 13. | Claim 20.....                                                        | 82 |
| 14. | Motivation to Combine.....                                           | 83 |
| C.  | Ground 3: Obviousness in light of Smyth, Phillips, and Nunes.....    | 85 |
| 1.  | Claim 4.....                                                         | 85 |
| 2.  | Claim 16.....                                                        | 86 |
| D.  | Ground 4: Obviousness in light of Smyth, Phillips, and Kuperman..... | 88 |
| 1.  | Claim 6.....                                                         | 88 |
| E.  | Ground 5: Obviousness in light of Smyth, Phillips, and Marin.....    | 89 |
| 1.  | Claim 7.....                                                         | 89 |
| 2.  | Claim 13.....                                                        | 89 |
| F.  | Ground 6: Obviousness in light of Smyth, Phillips, and Bozorgi.....  | 89 |
| 1.  | Claim 8.....                                                         | 90 |
| 2.  | Claim 12.....                                                        | 91 |
| IX. | CONCLUSION.....                                                      | 92 |

1. I have been retained by KKD Law PLLC, as an independent technical expert in the *Inter Partes* Review between CrowdStrike, Inc. (“CrowdStrike”) and Skysong Innovations, LLC (“Skysong”) PTAB Case No. IPR2025-01170 involving U. S. Patent No. 11,892,897 (“897 patent”).

2. I am being compensated for my work as a technical expert at my customary rate. My compensation does not in any way depend on the outcome of this review, and I have no personal interest in the outcome of this review. My declaration is a technical analysis of the subject patent and the art cited by Petitioner. Nothing in my declaration is intended to offer any legal conclusion.

## **I. QUALIFICATIONS**

3. My qualifications as an expert are detailed in my curriculum vitae, which is attached as **Appendix A** to this report. Appendix A includes a list of all publications authored or co-authored by me in the previous 10 years.

4. A list of all other cases in which, during the previous four years, I testified as an expert at trial or by deposition, is provided as **Appendix B** to this report.

5. I received a Bachelor of Science degree in Computer Science, summa cum laude, from Arizona State University’s Computer Science and Engineering department.

6. I received a Master of Science degree in Information and Computer Science from the University of California at Irvine's Information and Computer Science department.

7. I received a Doctor of Philosophy degree from the University of California at Irvine's Information and Computer Science department. My dissertation was entitled "Architecture-Based Specification-Time Software Evolution."

8. For over thirty years, my research has focused on software systems; their design, development, maintenance, and evolution; their individual building blocks and overall architectures; mechanisms for their implementation and deployment; and their critical properties, including security, trust, robustness, reliability, maintainability, etc.

9. My Ph. D. dissertation yielded a new design style for the development and evolution of highly distributed software systems built out of independent components that exchange different types of data encoded in messages to accomplish specific goals. This research focused on the design principles behind those systems, the appropriate support for modeling the systems (both at the level of individual modules as well as their integrations into subsystems and entire systems), and the infrastructure for their implementation, deployment, optimization, and long-term evolution. The techniques I developed relied on the novel notion of software connectors as loci of interaction in a system. Connectors are first-class architectural

building blocks, equivalent to functional components, and manage all aspects of interaction, spanning communication, data compression, encryption, integrity, coordination, distribution, etc. This architectural style and the tools I developed to support it were applied across a wide range of domains, from traditional distributed systems, to mobile devices and applications, to embedded systems and autonomous robots.

10. My work has been funded by US government agencies (e.g., the National Science Foundation, National Institutes of Health, Defense Advanced Research Projects Agency, NASA, and Department of Homeland Security) and industrial organizations (e.g., Boeing, IBM, Northrop Grumman, Xerox, Google, and Bosch).

11. I am employed by the University of Southern California (“USC”) as a faculty member in the Computer Science Department and have been since January 1999. I am a tenured professor. Since July 2022, I have served as my Department’s Chair. Between January 2009 and January 2013, I served as the Director of USC’s Center for Systems and Software Engineering (USC-CSSE), a large research center that involved the work of a number of USC faculty, students, and staff as well as several dozen industry partners. Under my direction, USC-CSSE was involved in a range of cutting-edge projects for our industry partners, ranging from the Aerospace Corporation to Xerox. Between July 2011 and July 2015, I served as my Department’s Associate Chair for Ph.D. Affairs. Between July 2021 and June 2022,

I served as my Department's Vice Chair for Tenure Affairs. In July 2025, I was named TRW Professor in Engineering and granted an endowed chair for my contributions to my university and the academic community.

12. In the mid-2000s, I worked with colleagues from NASA's Jet Propulsion Laboratory on designing the architecture of their OODT data grid system that was intended to provide the massive amounts of NASA's available planetary science data efficiently and securely to arbitrary numbers of users around the country. This work was jointly published at the 2006 International Conference on Software Engineering. Subsequently, I spearheaded a project at my university that extended OODT to make the data grid accessible to mobile devices.

13. Based on my experience and expertise in the area of software system architectures, and the success of the above project, in 2007 I was offered the position of Chief Software Engineer of the Transformational Satellite Communications System (TSAT) program. This was a \$4 billion U.S. Department of Defense (DOD) program, sponsored by the U.S. Air Force. The program's charge was to build the satellite grid of the future. With my university's agreement, I was prepared to go on a multi-year leave-of-absence assignment until the TSAT project was operational. Unfortunately, the program's funding was cut before my involvement could begin.

14. I have created and/or taught undergraduate-, Master's-, and PhD-level courses in software engineering, software requirements, software architecture, and

embedded systems for many years. As part of these courses, I have developed and taught modules, and created and guided projects that exposed students to the interplay between software and traditional computing devices on one hand, and a wide range of physical equipment on the other hand, including sensors, RFID systems and their integration into item tagging applications, cameras, mobile robots, and wearable computing devices.

15. I co-authored “*Software Architecture: Foundations, Theory, and Practice*,” a widely used textbook on software systems architecture initially published in 2009. The issues of security and trust are discussed throughout the book as guiding principles behind the design of modern software systems. Additionally, an entire chapter of the book is explicitly dedicated to this topic.

16. A critical aspect of all systems I have researched over the past three decades is system security. I have taught existing and developed new techniques for ensuring system security in distributed, mobile, and embedded settings. These techniques have spanned static program code analysis; dynamic monitoring, logging, simulation, emulation, and testing; and statistical analysis and algorithmic prediction based on historical data gathered both at the architectural/design and implementation levels.

17. In the early 2000s, I secured a grant from Intel Corp. which I used to equip a teaching laboratory with a variety of embeddable computing equipment. These are

non-traditional computing platforms, such as those listed above, that can be a part of devices used in everyday activities (e.g., “smart” cards, cameras, phones, clothing accessories, etc.). This laboratory was used as the foundation of a graduate-level course, Software Engineering for Embedded Systems, that I introduced soon thereafter and taught a number of times. This course covered a range of topics relevant to embedded systems, including efficient, reliable, and secure communication in networks of varying sizes and the challenges inherent therein. Every semester, my students had to complete a course project they selected from a range of real-world scenarios. These projects included real-time image tracking and recognition; incremental automated scene discovery and rendering; efficient and secure off-board data compression, storage, and access; and so on. Several of these projects grew into peer-reviewed publications and doctoral dissertations. Subsequent extensions of the projects from this course have been used externally, at NASA’s Jet Propulsion Laboratory and Bosch Research and Technology Center.

18. In 2007, I published a technique that was subsequently patented, in which I showed how software computation can be accomplished in a highly distributed manner on a public network such that mathematically provable guarantees are provided that no compromised nodes can discover the involved algorithm or the data (beyond the jumble of the binary bits involved).

19. Some of my earliest work relating to artificial intelligence (AI) pertained to determining the optimal deployment of a software-intensive system on a set of hardware devices. This project took started around 2000, was initially published in 2002, and culminated in two highly cited publications in the IEEE Transactions on Software Engineering journal, in 2005 and 2011. Initial attempts at solving the problem by taking into account large numbers of competing system requirements and device capabilities involved traditional (e.g., greedy) and linear programming algorithms. Those proved adequate, but had limitations. My students and I quickly turned to AI and adopted a genetic algorithm, which we further tailored in a number of innovative ways to outperform state-of-the-art techniques in this area.

20. In the mid 2010s, my research further explored the emerging AI and machine learning (ML) algorithms as vehicles for predicting various aspects of a software system. Initially, my lab developed algorithms that were applied to the predicting appropriate architectural design decisions that can resolve the implementation problems reported about a system. Subsequently, my students and I developed a pair of techniques, published in 2017 and 2018 and patented soon thereafter, where we analyze communications on a mobile device, identify potentially vulnerable interactions, and use a combination of static code analysis, dynamic/runtime analysis, and AI/ML algorithms to predict likely upcoming user

interactions with their device as a way of anticipating and blocking those interactions that may compromise the device's security.

21. My most recent work has focused on building a framework for evaluating AI/ML techniques—especially those relying on deep neural networks (DNNs)—that purport to pinpoint the sources of security vulnerabilities in software systems. My team's most recent publication on this topic, at the 2025 International Conference on Software Engineering, presents a framework called VIPer, which identified key shortcomings in existing DNN-based vulnerability detectors by showing how they learn from spurious system features and can be “fooled” by amplifying such features in their training datasets. VIPer provides an alternative neural network-based solution that significantly outperforms the current state-of-the-art in detecting vulnerabilities in previously unseen software systems.

22. A significant portion of my professional activity centers on conducting research through advising Ph.D. students. I have graduated 21 Ph.D. students to date, and I advise seven students currently pursuing their Ph.D. My students have gone on to successful careers in top software companies (e.g., Google, NVIDIA, Intel, Snap), research institutes (e.g., NASA's Jet Propulsion Laboratory, Information Sciences Institute, Quandary Peak Research), and universities (e.g., University of Massachusetts, University of California, University of Edinburgh in the UK).

23. I served as Program Co-Chair for the flagship conference in my field—International Conference on Software Engineering (“ICSE”)—held in May 2011. I have served as Chair or Co-Chair for various other conferences in the Software Engineering field, including: the Fifth Working Institute of Electrical and Electronics Engineers (“IEEE”)/IFIP Conference on Software Architecture, the Third IEEE International Conference on Self-Adaptive and Self-Organizing Systems, the Fifteenth International Association for Computing Machinery’s (“ACM”) Special Interest Group on Software Engineering (“SIGSOFT”) Symposium on Component Based Software Engineering, the IEEE/CSSE/ISE Workshop on Software Architecture Challenges for the 21st Century, and the Doctoral Symposia at the Sixteenth ACM SIGSOFT International Symposium on the Foundations of Software Engineering and at the 44th International Conference on Software Engineering.

24. I have served as an editor of several peer-reviewed journals, including IEEE Transactions on Software Engineering, ACM Transactions on Software Engineering and Methodology, Journal of Software Engineering for Robotics, Elsevier Information and Software Technology Journal, Journal of Systems and Software, Journal of Software Engineering Research and Development, and Springer Computing Journal. I have also served as a guest editor of several special issues for other journals. Between January 2018 and December 2022, I served as

Editor-in-Chief of IEEE Transactions on Software Engineering (“TSE”), the flagship software engineering research journal.

25. Between September 2013 and September 2015, I served as Chair of the ICSE Steering Committee. Both before that (2008-2013) and subsequent to that (2015-present) I have been a member of the ICSE Steering committee. I am currently a member of the Steering Committee of the joint European Conference on Software Engineering and the Symposium on the Foundations of Software Engineering (“ESEC/FSE”). From July 2015 to September 2017, I chaired the Steering Committee of FSE, while it was a standalone symposium. I have served as a member of the Steering Committees of the Working IEEE/IFIP Conference on Software Architecture on two separate occasions, including presently.

26. Between July 2015 and June 2018, I served as Chair of the ACM SIGSOFT, the largest professional organization in my field. During the past year, I have chaired the committee in charge of ACM’s Software System Award, one of the most important technical awards in my field.

27. I have served as editor of various books in the Software Engineering field, including Proceedings of the 3rd International Conference on Self-Adaptive and Self-Organizing Systems, Proceedings of the Warm-Up Workshop for the 32nd International Conference on Software Engineering, Proceedings of the 5th Working

IEEE/IFIP Conference on Software Architecture, and 40 Editions of ICSE—the Ruby Anniversary Celebration.

28. I have authored or co-authored nearly 300 papers in the Software Engineering field. My work has been cited over 21,000 times. My most cited paper alone has been cited nearly 3,200 times; that paper recently received a Most Influential Paper Award from IEEE TSE. A paper I co-authored in the 1998 ICSE, my field’s flagship conference, was given ten years later, in 2008, that conference’s Most Influential Paper Award. Another of my publications received a Most Influential Paper Award in 2020. Several of my other publications have received “best paper” or “most cited” awards.

29. My research frequently results in the development of software tools that accompany and demonstrate the published ideas. The above-mentioned paper describing a security-vulnerability detection and remediation tool for mobile (e.g., Android phone) devices built in my lab, won the Best Tool Paper award at the 2017 Automated Software Engineering Conference. That idea was subsequently awarded a US patent. I am a co-inventor on two additional issued patents in distributed systems design, implementation, deployment, and computer security.

30. I have been named a Distinguished Scientist of the ACM. I have been elected a Fellow of the IEEE. Fellow of the IEEE is IEEE’s highest grade; less than 0.1% of its membership earns this grade each year.

## II. BASES OF OPINIONS

31. The basis and reasoning of my opinions include my education, training, and experience as an engineer, including my decades of experience in computer science. In conducting my analysis and forming my opinions, I have considered the materials listed below:

- a. U.S. Patent No. 11,892,897 to Paulo Shakarian, *et al.*, filed Oct. 26, 2018, and issued on Feb. 6, 2024, its provisional applications, and its file history (EX1002);
- b. CrowdStrike's petition (Paper 1, "Pet.") and any materials referenced therein relevant to my analysis; specifically including the declaration of Dr. Nicholas Bambos (EX1003, "Bambos");
- c. Carl Sabottke et al., *Vulnerability Disclosure in the Age of Social Media: Exploiting Twitter for Predicting Real-World Exploits*, Proceedings of the 24th USENIX Security Symposium (2015) (EX1004, "Sabottke")
- d. U.S. Patent No. 10,264,009 to Smyth et al. (EX1005, "Smyth")
- e. Elizabeth Phillips et al., *Extracting Social Structure from DarkWeb Forums*, SOTICS 2015: The Fifth International Conference on Social Media Technologies, Communication, and Informatics, (2015) (EX1006, "Phillips")

- f. Eric Nunes et al., *Darknet and Deepnet Mining for Proactive Cybersecurity Threat Intelligence*, Proceedings of the 2016 IEEE Conference on Intelligence and Security Informatics (EX1007, “Nunes”)
- g. U.S. Patent No. 10,652,254 to Kuperman et al. (EX1008, “Kuperman”)
- h. Ericsson Marin et al., *Product Offerings in Malicious Hacker Markets*, Proceedings of the 2016 IEEE Conference on Intelligence and Security Informatics (EX1009, “Marin”)
- i. Mehran Bozorgi et al., *Beyond Heuristics: Learning to Classify Vulnerabilities and Predict Exploits*, Proceedings of the Sixteenth ACM Conference on Knowledge Discovery and Data Mining (KDD 2010) (EX1010, “Bozorgi”)
- j. and any other materials referenced herein.

### **III. A PERSON OF ORDINARY SKILL IN THE ART**

32. My opinions in this declaration are based on the understanding of a Person of Ordinary Skill in the Art (“POSITA”) at the time of the invention of the claims in the ’897 patent.

33. In assessing the level of skill of a POSITA, I have considered the type of problems encountered in the art, the prior solutions to those problems found in the prior art references, the rapidity with which innovations are made, the sophistication

of the technology, the level of education of active workers in the field, and my own experience working with those of skill in the art at the time of the invention.

34. In my opinion, a person of ordinary skill in the art at the time of the filing of the '897 patent would have at least a Bachelor's degree in CS, data science/engineering, or cybersecurity, and 2–4 years of experience working in the field of software development, AI, ML, data mining, or cybersecurity. A graduate degree in CS, AI, data science/engineering, cybersecurity, or software engineering may be substituted wholly or in part for work experience.

35. I am very familiar with this level of skill. Over my decades of experience, I have supervised and worked with engineers in this field having at least the level of skill identified above.

#### **IV. LEGAL FRAMEWORK**

##### ***A. Claim Construction***

36. I understand that the words of a claim are generally given their ordinary and customary meaning, that is, the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention.

37. I understand that a claim term that does not use the word “means” is presumed not to be a means-plus-function term. I understand that a term that does not use the word “means” would be construed as a means-plus-function term if it

does not describe structure to a POSITA. Conversely, I understand that if a term describes structure to a POSITA, it is not a means-plus-function term.

38. I understand that a patent has several components, including an abstract, drawings, a written description detailing different embodiments of the invention (*i.e.*, the specification), and numbered claims at the end of the patent. It is the numbered claims that define metes and bounds of the properties that the patentee has the right to exclude other from infringing.

***B. Validity***

39. I understand that a patent, once issued or granted, is entitled to a presumption of validity. However, this presumption can be overcome by clear and convincing evidence to the contrary. Validity is routinely challenged post-issuance, and courts often invalidate the patent, on the ground that the patent fails to meet one or more statutory requirements, *e.g.*, by claiming a non-patentable subject matter, such as an abstract idea, in violation of 35 U.S.C. § 101; by being not novel and anticipated by prior art under § 102; by being obvious in view of prior art under § 103; and/or lacking sufficient disclosures or being indefinite, in violation of § 112.

40. I understand that the presumption of validity does not apply before the PTAB.

## **1. Anticipation**

41. I understand that section 102 first defines “prior art” that would anticipate and, thus, invalidate the patent. The term “prior art” is a patent term of art for the technology that was known or used before the priority date of the patent, which is typically the filing date of the patent application. However, I understand that, to the extent a provisional patent application properly discloses and supports the claimed subject matter, the priority date of the claims is the filing date of the provisional application, which is considered the invention date. Thus, if the subject matter of the invention was already patented or described in a printed publication anywhere in the world before the invention claimed in the patent, such prior art would invalidate the patent. 35 U.S.C. § 102(a).

42. Anticipation is a technical defense. A patent claim is anticipated and therefore invalid under 35 U.S.C. §102, only if: a single prior art reference discloses, either expressly or inherently, each and every element of the claimed invention arranged exactly as they are arranged in the claim. If even a single element or step recited by the claim is not disclosed in exactly the same way, arranged as in the claim, there is no anticipation. In addition, the reference must enable one of skill in the art to practice an embodiment of the claimed invention without undue experimentation.

43. For an element to be present in a reference through inherency, the evidence must be clear and convincing that the missing element is necessarily present in the thing described in the reference, and that it would be so recognized by persons of ordinary skill. That means that inherency cannot be established by probabilities or possibilities. The mere fact that a certain thing may result from a given set of circumstances is not sufficient. If, however, the disclosure is sufficient to show that the natural result flowing from the operation as taught would result in the performance of the questioned function, that disclosure could be inherent.

44. I understand that CrowdStrike does not argue that any prior art anticipates the '897 patent.

## **2. Obviousness**

45. I understand that § 103 is broader than § 102 in a sense that § 103 forbids issuance of a patent even if the invention is not identically disclosed or described as set forth in section 102, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.

46. I understand that to measure obviousness under § 103, a court determines the scope and content of the prior art, ascertains differences between the prior art and the claims at issue, and resolves the level of ordinary skill in the pertinent art. I

understand that secondary considerations such as commercial success, long felt but unsolved needs, failed attempts or copying by others, unexpected results, professional praise, licensing activity, expert surprise, or departing from prior-art teachings, may illuminate the circumstances surrounding the invention's origin.

47. It is my understanding that the information that is used to evaluate whether an invention is new and not obvious is generally referred to as "prior art." Prior art generally includes U.S. and foreign patents and U.S. and foreign printed publications (e.g., published patent applications (before issued as a patent), books, journal publications, presentation files, posters, articles on websites, product manuals, etc.) that existed before the earliest filing date (the "effective filing date") of the claims in the patent. A patent will be prior art if it was filed before the effective filing date of the claimed invention, while a printed publication will be prior art if it was publicly available before that date.

48. I understand that, in addition to the patents and printed publications, a product used in public or on sale in the U.S. more than one year prior to the effective filing date of a patent can be prior art to that patent. I understand that a product is considered "on sale" when (i) the product embodying the invention is offered for commercial sale, and (ii) the invention was ready for patenting. An invention is shown to be "ready for patenting" when there is proof of a reduction to practice or proof that prior to the critical date the inventor had prepared drawings or other descriptions of the invention that were sufficiently specific to enable a person skilled

in the art to practice the invention. I understand that CrowdStrike does not assert a product as prior art.

49. It is my understanding that a claimed invention is not patentable if it would have been obvious to a person of ordinary skill in the field of the invention at the time the invention was made, and that the burden of establishing obviousness rests with the challenger and must be supported by clear and convincing evidence.

50. It is my understanding that the following standards govern the determination of whether a claim in a patent is obvious. I have applied these standards in my evaluation, recognizing that obviousness determinations must be based on a careful evidence-based analysis rather than general assertions or hindsight reasoning.

51. It is my understanding that, to find a claim in a patent obvious, one must make certain findings regarding the claimed invention and the prior art. Specifically, the obviousness question requires consideration of four factors (although not necessarily in order):

- a. the scope and content of the prior art;
- b. the differences between the prior art and the claims at issue;
- c. the knowledge of a person of ordinary skill in the pertinent art; and

- d. objective factors that may indicate non-obviousness, which must be given meaningful weight rather than treated as secondary or optional considerations.

52. It is my understanding that the obviousness inquiry should not be done in hindsight, but must be done using the perspective of a person of ordinary skill in the relevant art as of the effective filing date of the patent claim. This includes avoiding reconstruction of the invention with knowledge of the patent in hand, and avoiding attributing to the prior art recognition of the very problem solved by the invention.

53. It is my understanding that the obviousness inquiry may also consider certain objective indicia of non-obviousness, which serve as a critical safeguard against hindsight. Such objective factors indicating non-obviousness may include: commercial success of products covered by the patent claims; a long-felt need for the invention; failed attempts by others to make the invention; copying of the invention by others in the field; unexpected results achieved by the invention; praise of the invention by the infringer or others in the field; the taking of licenses under the patent by others; expressions of surprise by experts and those skilled in the art at the making of the invention; and the patentee proceeded contrary to the accepted wisdom of the prior art. These considerations are not merely supplemental, but can be dispositive.

54. It is my understanding that the fact finder must determine whether potential evidence of such objective considerations is relevant. With respect to evidence offered for each objective consideration, the fact finders must ascertain whether there is a nexus between the claimed invention and the evidence establishing the objective consideration and determine its probative value. Where such a nexus is present, the evidence must be given substantial weight in favor of non-obviousness.

55. I understand that the obviousness analysis requires a comparison of the properly construed claim language to the prior art on a limitation-by-limitation basis. The analysis may not rely on broad generalizations or assumptions about what a person of ordinary skill might have done. Only if a person of ordinary skill in the art can implement the claimed invention as a predictable variation of a known invention without undue experimentation and with a clear motivation grounded in the prior art may a claim be considered obvious. Also, “hindsight” reconstruction cannot be used to combine references together to reach a conclusion of obviousness.

56. It is my understanding that exemplary rationales that may support a conclusion of obviousness include:

- a. combining prior art elements according to known methods to yield predictable results;
- b. simple substitution of one known element for another to obtain predictable results;

- c. use of known techniques to improve similar devices (methods or products) in the same way;
- d. applying a known technique to a known device (method or product) ready for improvement to yield predictable results;
- e. “obvious to try,” *i.e.*, choosing from a finite number of identified, predictable solutions but only where there is a demonstrable and reasonable expectation of success at the time;
- f. known work in one field may prompt variations for use in another field only where such variations would have been reasonably predictable and supported by the record; and
- g. some teaching, suggestion, or motivation in the prior art that is contemporaneous and specific rather than assumed through hindsight.

57. Thus, I understand that a reference may be relied upon for all that it teaches, including uses beyond its primary purpose. However, where a reference teaches away, *i.e.*, discourages or criticizes the path taken by the invention, such teachings must be given meaningful weight. The mere disclosure of alternative designs does not teach away, but explicit indications that a path would be undesirable, inferior, or impractical strongly weigh against obviousness.

58. I understand that a claim might be obvious over a single reference; however, such an argument requires that the reference clearly and fully provide the

missing teachings or that such teachings would have been immediately apparent to a person of ordinary skill without inventive insight, not based on generalized knowledge or hindsight interpretation.

59. An obviousness evaluation can also be based on a combination of multiple prior art references. However, a specific and articulated rationale must exist to justify such a combination, and the motivation must be grounded in the prior art itself, not reconstructed from the claim. I further understand that, although market demand may influence innovation, it cannot substitute for a teaching, suggestion, or motivation in the prior art.

60. I understand that a proper obviousness analysis must recognize that innovation often results from insight that was neither expected nor recognizable from merely assembling known pieces in the prior art. A person of ordinary skill in the art must be considered without the benefit of hindsight, and thus cannot be assumed to have known the solution or outcome in advance. Therefore, any alleged combination of prior art must be shown to have been clearly supported and reasonably suggested by the prior art itself at the relevant time, and not inferred only after knowledge of the invention's success.

61. I understand that the concept of "obvious to try" applies only in rare circumstances, and only where the possible options were finite, predictable, and accompanied by a reasonable expectation of success. If experimentation, insight, or

creativity were required to arrive at the claimed invention, the invention is not obvious.

62. It is my understanding that combining familiar elements using known methods does not render a claim obvious where the result yields new, unexpected, or improved properties or solves a previously unrecognized problem. Obviousness requires predictability, and the absence of such predictability strongly indicates non-obviousness.

63. I understand that when a patent arranges known elements, the question is whether the arrangement yields more than predictable results and reflects inventive judgment. The analysis focuses on what was known or knowable to a person of ordinary skill at the time, and whether the problem addressed was recognized and appreciated in the prior art, not merely whether solutions appear straightforward after the fact.

64. In sum, it is my understanding that prior art teachings may be combined only where the record demonstrates that a person of ordinary skill, without hindsight and based on contemporaneous knowledge, would have been motivated to do so with a reasonable expectation of success. Where such evidence is lacking, or where hindsight reconstruction is required, obviousness may not be found.

## V. THE '897 PATENT (EX1001)

### A. *Technical Background*

65. My below discussion pertains to the state of the art at the time of the invention of the '897 patent unless specifically indicated otherwise.

66. The '897 patent teaches advanced artificial intelligence algorithms for assessing the likelihood of exploitation of software vulnerabilities by malicious hackers and predicting which software vulnerabilities are likely to be exploited first to prioritize patching. '897 patent at abstract and 1:24–28.

67. The '897 patent first describes the known problem, i.e., the ever increasing number of software vulnerabilities that are discovered and publicly disclosed every year. '897 patent at 1:32–33. “In 2016 alone, more than 10,000 vulnerability identifiers were assigned and at least 6,000 were publicly disclosed by the National Institute of Standards and Technology (NIST).” *Id.* at 1:33–36. “Once the vulnerabilities are disclosed publicly, the likelihood of those vulnerabilities being exploited increases.” *Id.* at 1:36–38. Indeed, “Verizon reported that over 99% of breaches are caused by exploits to known vulnerabilities.” *Id.* at 1:49–50.

68. The time to react to the publication of these vulnerabilities is often critically short. *Id.* at 1:51–53. “For instance, exploits targeting the Heartbleed bug in OpenSSL were detected in the wild 21 hours after the vulnerability was publicly disclosed.” *Id.* at 1:53–57. The problem is exacerbated by the sheer number of

known vulnerabilities, only ~3% of which may actually be exploited. *Id.* at 1:62–64; *see also id.* at 4:64–65 (“around 2.4% of the reported vulnerabilities are exploited in real world attacks.”); *see also id.*, table 3 (vulnerability exploitation probabilities); *id.* at 13:50–51 (referencing “notable variations on the exploitation likelihood”).

69. For these reasons, there is a critical need “to efficiently assess vulnerabilities that will be exploited in the wild while keeping the false alarm rate low.” *Id.* at 1:57–59.

70. The known prior art has many shortcomings. For example, standard risk assessment systems provided by software makers, such as “Common Vulnerability Scoring System (CVSS), Microsoft Exploitability Index, and Adobe Priority Rating are overly cautious and report “many vulnerabilities as ‘severe’ and ‘will be exploited’.” *Id.* at 1:41–45. However, since “only a small fraction of those vulnerabilities (less than 3%) are found to be exploited in the wild” (*id.* at 1:62–64), these identifications do “not alleviate the problem much since the majority of the flagged vulnerabilities will not be attacked.” *Id.* at 1:45–47. “Further, previous work has found that the CVSS score provided by NIST is not an effective predictor of vulnerabilities being exploited.” *Id.* at 1: 63–67.

71. Other prior art has explored “the use of social media, darkweb markets and certain white-hat websites” to attempt to prioritize vulnerabilities. *Id.* at 1:67–2:2. However, “methodical concerns on the use of social media for exploit prediction

were recently raised; data feeds for proactive exploit prediction were limited to single sites that provided a relatively small number of predictions.” *Id.* at 2:4–7. Notably, other work has “demonstrated the viability of data collection but does not quantify the results of prediction.” *Id.* at 2:8–9.

72. The invention of the ’897 patent directly addresses these problems. Specifically, the ’897 patent examines the ability to predict exploits in the wild “using cyber threat intelligence feeds comprised from a variety of data sources or data feeds.” *Id.* at 3:27–29. This is “of direct relevance to patch prioritization.” *Id.* at 3:30.

73. One specific important advantage of the ’897 patent is its ability to analyze and rely on a variety of data sources or data feeds. “Many sources of data were identified that were representative of current threat intelligence used for vulnerability prioritization-examples being presented in FIG. 13.” *Id.* at 3:8–41, Fig. 13.

74. For example, the ’897 patent discloses “Embodiment A” which examines three sources (ExploitDB (EDB), Zero Day Initiative (ZDI) by TippingPoint, and a collection of information scraped from over 120 sites on the darkweb and deepweb (DW)). *Id.* at 3:42–50. The second embodiment “Embodiment B” utilizes “a database of 46 darkweb and deepweb forums collected from the same data collection

infrastructure used by embodiment A,” but over a longer time period of data. *Id.* at 9:50–55.

75. In summary, the ’897 patent teaches training a machine learning model where the training data consists of multiple datasets (going beyond NVD) on vulnerability information, and multiple datasets on hacker communications. (The ’897 patent describes the use of over 40 hacker websites. ’897 patent at 9:50–53. The features in this dataset include explicit data, metadata, and social connections. Further, the ’897 patent uses special measures of the social connections relying on ideas of social science theory. The patent notes that the model is trained on data that occurred prior to when the vulnerability was exploited, and there is also data prior to the disclosure date. The target of the prediction is well-defined as evidence of exploits that occur in the wild, as opposed to exploits that are developed as “proof of concept”. The ’897 patent further uses proof-of-concept exploit information to predict the exploits in the wild, and it teaches multiple different classifications of likelihood of exploitation, such as “not likely”, “likely”, “extremely likely”.

### ***B. Challenged Claims***

76. CrowdStrike’s Petition challenges all claims (1–20) of the of the ’897 patent. Of these, claims 1, 14, and 18 are independent.

## **1. Claim 1**

**Claim 1.** A method for assessing a likelihood of exploitation of software vulnerabilities, comprising:

- a.** utilizing a processor in operable communication with at least one memory for storing instructions that are executed by the processor to perform operations, including:
- b.** accessing a plurality of datasets associated with a predetermined set of data sources, the plurality of datasets including training data comprising hacker communications;
- c.** accessing features from the plurality of datasets that include measures computed from social connections of users posting hacking-related content
- d.** applying learning algorithms to the training data to generate classification models that are configured to predict class labels defining a likelihood of exploitation of respective software vulnerabilities;
- e.** accessing one or more features associated with a software vulnerability; and
- f.** computing, by applying the one or more features to the classification model, a class label defining one or more values defining a likelihood of exploitation associated with the software vulnerability,
- f.1.** wherein the likelihood of exploitation predicts an actual exploitation of the respective software vulnerabilities before disclosure based on the hacker communications from the training data.

## **2. Claim 2**

**Claim 2.** The method of claim 1, further comprising generating a plurality of estimation outputs based on the one or more values to derive an overall quantitative score.

## **3. Claim 3**

**Claim 3.** The method of claim 1, wherein the plurality of datasets include vulnerability data for vulnerabilities that are publicly disclosed and obtaining exploits data for exploits that were used in real world attacks.

#### **4. Claim 4**

**Claim 4.** The method of claim 3, further comprising:

aligning the exploits data with the vulnerability data; and

cleaning the exploits data of noise and predetermined portions of the exploits data that is irrelevant to associated software vulnerabilities.

#### **5. Claim 5**

**Claim 5.** The method of claim 1, wherein certain features correspond to a known vulnerability obtained from the plurality of datasets.

#### **6. Claim 6**

**Claim 6.** The method of claim 1, further comprising testing the classification models by applying additional training data and one or more algorithms and evaluation metrics to optimize the classification models until the classification models compute the likelihood of exploitation according to a predefined error rate.

#### **7. Claim 7**

**Claim 7.** The method of claim 1, further comprising vectorizing text features derived from the plurality of datasets using term frequency-inverse document frequency to create a vocabulary of associated words.

#### **8. Claim 8**

**Claim 8.** The method of claim 1, further comprising:

sorting vulnerabilities associated with the plurality of datasets according to time;

training the classification model using the training data, the training data defining a first subset of the plurality of datasets associated with a predetermined period of time; and

testing the classification model using a second subset of the plurality of datasets associated with the predetermined period of time.

## **9. Claim 9**

**Claim 9.** The method of claim 1, further comprising computing mutual information from the plurality of datasets informative as to what information a given feature provides about another feature.

## **10. Claim 10**

**Claim 10.** The method of claim 1, further comprising:

detecting, from the plurality of datasets, vulnerabilities that appear before an associated exploitation date.

## **11. Claim 11**

**Claim 11.** The method of claim 1, further comprising:

accessing features from the plurality of datasets that measure a centrality of the users in a social graph.

## **12. Claim 12**

**Claim 12.** The method of claim 1, further comprising:

accessing one or more features indicative of temporal connections between at least two of:

- a time associated with discussion of a vulnerability by users posting hacking-related content at a web forum prior to disclosure to a public vulnerability database;

- a time associated with disclosure of the vulnerability to the public vulnerability database; and

- a time associated with exploitation of the vulnerability as obtained through exploits data associated with real-world exploitation of the vulnerability.

## **13. Claim 13**

**Claim 13.** The method of claim 1, further comprising vectorizing text features derived from textual content of the plurality of datasets using a predetermined natural language process (NLP [sic]).

#### **14. Claim 14**

**Claim 14.** A computing device, configured via machine learning to apply a learned function to data associated with a software vulnerability to estimate a likelihood of exploitation of the software vulnerability, the learned function associated with a prediction model derived from at least one machine learning algorithm and a plurality of datasets associated with software vulnerabilities, the plurality of datasets including information associated with discussion of vulnerabilities by users posting hacking-related content, wherein the likelihood of exploitation predicts an actual exploitation of the software vulnerability before disclosure based on hacker communications from the plurality of datasets.

#### **15. Claim 15**

**Claim 15.** The computing device of claim 14, wherein the prediction model is at least one classification model that outputs from features of the software vulnerability a score indicative of the likelihood of exploitation.

#### **16. Claim 16**

**Claim 16.** The computing device of claim 14, wherein the predictive model includes a Random Forest (RF) method including multiple decision tree predictors applied in combination to classify the software vulnerability which is used to estimate the likelihood of exploitation.

#### **17. Claim 17**

**Claim 17.** The computing device of claim 14, wherein the computing device is further configured to apply the learned function to data associated with each of a plurality of new software vulnerabilities and compute a respective likelihood of exploitation for each of the plurality of new software vulnerabilities.

#### **18. Claim 18**

**Claim 18.** A method of prioritizing vulnerabilities using cyber threat intelligence, comprising:

**a.** utilizing a processor in operable communication with at least one memory for storing instructions that are executed by the processor to perform operations including:

**b.** accessing a plurality of datasets associated with a predetermined set of data sources, at least a portion of the plurality of datasets defining training data including hacker communications;

**c.** applying learning algorithms to the training data to generate a predictive model configured to predict a likelihood of exploitation of respective software vulnerabilities;

**d.** accessing one or more features associated with a software vulnerability;  
and

**e.** computing, by applying the one or more features to the predictive model, one or more values defining a likelihood of exploitation associated with the software vulnerability, wherein the likelihood of exploitation predicts an actual exploitation of the respective software vulnerabilities before disclosure based on the hacker communications from the training data.

#### **19. Claim 19**

**Claim 19.** The method of claim 18, further comprising generating a plurality of estimation outputs based on the one or more values to derive an overall quantitative score.

#### **20. Claim 20**

**Claim 20.** The method of claim 18, wherein the plurality of datasets include vulnerability data for vulnerabilities that are publicly disclosed and obtaining exploits data for exploits that were used in real world attacks.

#### ***C. Priority Dates***

77. The '897 patent derives priority from provisional application No. 62/581,123, filed on Nov. 3, 2017. I have specifically reviewed the provisional application to confirm that the subject matter claimed in the '897 patent is disclosed therein. Although it is my opinion that the claimed subject matter is in fact disclosed in the original provisional patent applications, I understand that Petitioner alleges

the following priority dates for the asserted prior art, which predate the provisional application:

- a. *Sabottke* (EX1004): September 23, 2015 (Pet. at 5)
- b. *Smyth* (EX1005): July 26, 2016 (Pet. at 8)
- c. *Phillips* (EX1006): August 3, 2016 (Pet. at 10)
- d. *Nunes* (EX1007): November 8, 2016 (Pet. at 11–12)
- e. *Kuperman* (EX1008): Feb. 23, 2017 (Pet. at 13)
- f. *Marin* (EX1009): November 8, 2016 (Pet. at 14)
- g. *Bozorgi* (EX1010): February 28, 2014 (Pet. at 15)

78. For my analysis in this declaration only, and without forming an opinion as to their accuracy, I have accepted these priority dates. My technical analysis therefore relies on citations to the specification of the '897 patent. However, I reserve the right to update my analysis with citations to the relevant provisional application in the event that I am informed later that the priority dates asserted by Petitioner are not accurate and in fact post-date the provisional patent application.

## **VI. THE ASSERTED PRIOR ART**

### ***A. Sabottke (EX1004)***

79. *Sabottke* refers to the paper, “Vulnerability Disclosure in the Age of Social Media: Exploiting Twitter for Predicting Real-World Exploits” by Carl *Sabottke, et al.* It explores the use of Twitter to predict which newly disclosed software

vulnerabilities are likely to be exploited in real-world attacks. EX1004 at 2. *Sabottke* addresses “need for prioritizing the response to new disclosures by assessing which vulnerabilities are likely to be exploited and by quickly ruling out the vulnerabilities that are not actually exploited in the real world.” *Id.* *Sabottke* notes that existing vulnerability scoring systems often err on the side of caution and that a better method is needed to prioritize responses to new disclosures. *Id.* Notably, *Sabottke* relies exclusively on Twitter for its input data.

80. *Sabottke* notes that “social media sites like Twitter, constitute rich sources of information, as the participants in coordinated disclosures discuss technical details about exploits and the victims of attacks share their experiences.” *Id.* at 2. *Sabottke* explores “the opportunities for early exploit detection using information available on Twitter,” and seeks to “characterize the exploit-related discourse on Twitter, the information posted before vulnerability disclosures, and the users who post this information.” *Id.*

81. *Sabottke* claims three contributions: (1) characterization of the landscape of threats related to information leaks about vulnerabilities before their public disclosure with automatic feature extraction from Twitter to detect exploits; (2) description of the first technique for early detection of real-world exploits using social media; and (3) introduction of a specific threat model and evaluation of the robustness against adversarial interference. *Id.* at 3.

82. *Sabottke* does not consider “the detection of zero day attacks (which exploit vulnerabilities before their public disclosures.” *Id.* at 3. *Sabottke* further explicitly notes that “a complete system for early exploit detection . . . likely requires mechanisms beyond the realm of Twitter analytics” “is also out of scope for this paper.”

83. *Sabottke* further explicitly identifies “data scarcity” as a significant challenge to its model.

**Data scarcity.** Prior research efforts on Twitter analytics have been able to extract information from millions of tweets, by focusing on popular topics like movies [27], flu outbreaks [20, 26], or large-scale threats like spam [56]. In contrast, only a small subset of Twitter users discuss vulnerability exploits (approximately 32,000 users), and they do not always mention the CVE<sup>1</sup> numbers in their tweets, which prevents us from identifying the vulnerability discussed. In consequence, 90% of the CVE numbers disclosed during our observation period appear in fewer than 50 tweets. Worse, when considering the known real-world exploits, close to half have fewer

---

<sup>1</sup> CVE: Common Vulnerabilities and Exposures (CVE) first published in the National Vulnerability Database (NVD). ’897 patent at 2:28–29.

than 50 associated tweets. This data scarcity compounds the challenge of class imbalance for reducing false positives and false negatives.

*Id.* at 4.

84. In summary, *Sabottke* trains a machine learning model where the training data is restricted to NVD, OSVDB, and Twitter data—wherein the Twitter data is the only dataset potentially including hacker data. This information is used to predict exploits that are developed—though not necessarily used—in an attack. *Sabottke* only considers vulnerabilities after the disclosure date. The social information on Twitter is used directly as features and is not further analyzed with other methods. Further, *Sabottke*'s training data includes information that occurred after exploitation in the wild, by using “stratified random samples.” Further, the target of prediction is defined as any exploit—whether used by hackers in an attack or not. Finally, *Sabottke* only yields one of two classes: “exploited” or “non-exploited.” No further indication of “likelihood of exploitation” is produced.

**B. *Smyth (EX1005)***

85. *Smyth* refers to U.S. Patent No. 10,264,009 to Smyth et al. (EX1005, “*Smyth*”), titled “*Automated Machine Learning Scheme for Software Exploit Prediction*.” Generally, *Smyth* discloses “a predictive engine for analyzing existing vulnerability information to determine the likelihood of a vulnerability being

exploited by malicious actors against a particular computer or network of computers.” *Phillips* at abstract.

86. More specifically, *Smyth* attempts to predict the likelihood that a current vulnerability may be “employed in a customer computer network.” *Smyth* utilizes a combined data set of historical vulnerability information, current vulnerability information, and a plurality of data sources. *Smyth* at 2:34–38.

87. *Smyth* describes “data sources” as repositories of information describing vulnerabilities that have been observed and characterized. *Smyth* at 4:36–37. These include, for example, NVD, “Virus Total Samples and Reports,” and “online discussions 240 including those on social media and discussion forums.” *Smyth* at 4:45–47. “A common characteristic of the data sources 210, 220, 230, 240, 250, 260 is that they rely on a common identifier associating information with a particular vulnerability.” *Smyth* at 4:52–55.

88. *Smyth* further notes that “CVE identifiers are also often used by security professionals and other interested individuals when discussing exploitation of vulnerabilities either through malware or through the acts of malicious actors. These discussions may occur in private and public contexts such as security-focused forums, blogs, websites, and social media posts.” 5:10–13.

**C. *Phillips (EX1006)***

89. *Phillips* refers to Elizabeth Phillips et al., *Extracting Social Structure from DarkWeb Forums*, SOTICS 2015: The Fifth International Conference on Social Media Technologies, Communication, and Informatics, (2015) (EX1006, “*Phillips*”).

90. *Phillips* focuses exclusively on the research question “Using communication metadata alone, can we make reasonable inferences about the structure of terrorist groups.” *Phillips* at 97. Within that paradigm, *Phillips* aims to “identify a range of potentially ‘important’ members of Islamic Networks within Dark Web Forums.” *Id.* Notably, “whilst existing work attempts to identify these structures through state-of-the-art Computational Linguistic techniques, [*Phillips*] relies on the communication metadata alone.” *Id.*

91. *Phillips* intentionally ignores communication content because “whilst existing research has been undertaken into creating a social network topology for Dark web Forums [11], these existing research focus on small communities and networks and rely on sophisticated sentiment analysis techniques, which are difficult to scale when evaluating millions of message.” *Id.*

**D. Nunes (EX1007)**

92. *Nunes* refers to Eric Nunes et al., *Darknet and Deepnet Mining for Proactive Cybersecurity Threat Intelligence*, Proceedings of the 2016 IEEE Conference on Intelligence and Security Informatics.

93. *Nunes* discloses a system for “cyber threat intelligence gathering from various social platforms on the Internet particularly sites on the darknet and deepnet.” *Nunes* at abstract. *Nunes* collects information “from hacker forum discussions and marketplaces offering products and services focusing on malicious hacking . . . for the purposes of identifying emerging cyber threats.” *Id.* Specifically, *Nunes* aims to identify “newly developed malware and exploits that have not yet been deployed in a cyber-attack.” *Id.*

94. *Nunes* discloses an “implementation and evaluation of learning models to separate relevant information from noise in the data collected from these online platforms.” *Id.* at 2. Notably, *Nunes* does not aim to determine the likelihood of exploitation of a given vulnerability. It’s sole focus is on identifying products (from market data) and communications (from forum data) relevant to potential threats. *Nunes* at 2.

**E. Kuperman (EX1008)**

95. *Kuperman* refers to U.S. Patent No. 10,652,254 to Kuperman et al. (EX1008, “*Kuperman*”), titled “*Analyzing Web Application Behavior To Detect*

*Malicious Requests.*” Generally, *Kuperman* discloses a system for “protecting web applications at a host by analyzing web application behavior to detect malicious client requests.” *Kuperman* at abstract.

96. More specifically, *Kuperman* discloses a proxy between client and server which includes first a list of known clients (malicious and non-malicious), and second a web application firewall (WAF) for classifying unknown requests. *Id.* The WAF generates a model based on the attributes of known malicious and non-malicious requests based on their attributes. *Id.*

97. *Kuperman* discloses that the WAF may utilize an Artificial Neural Network (ANN) which is trained and may be tested and retrained in order to achieve a desired error rate. *Kuperman*, Fig. 5 and related description (18:2–22:7).

***F. Marin (EX1009)***

98. *Marin* refers to Ericsson Marin et al., *Product Offerings in Malicious Hacker Markets*, Proceedings of the 2016 IEEE Conference on Intelligence and Security Informatics (“*Marin*”).

99. *Marin* discloses scraping darkweb and deepweb marketplaces for malicious hacking products in order to collect information about such products in a unified database schema. *Marin* at abstract. More specifically, *Marin* teaches collecting and analyzing marketplace data using “standard term frequency - inverse document frequency (TF-IDF) vectorization. *Marin* at 2.

**G. *Bozorgi (EX1010)***

100. *Bozorgi* refers to Mehran Bozorgi et al., *Beyond Heuristics: Learning to Classify Vulnerabilities and Predict Exploits*, Proceedings of the Sixteenth ACM Conference on Knowledge Discovery and Data Mining (KDD 2010) (“*Bozorgi*”).

101. *Bozorgi* discloses a system to determine whether and how soon “individual vulnerabilities are likely to be exploited.” *Bozorgi* at abstract. Specifically, *Bozorgi* relies on publicly available databases that document past vulnerabilities, to “show how to train classifiers that predict whether and how soon a vulnerability is likely to be exploited.” *Id.* “As input, [*Bozorgi*’s] classifiers operate on high dimensional feature vectors that we extract from the text fields, time stamps, cross-references, and other entries in existing vulnerability disclosure reports.” *Id.*

**VII. CLAIM CONSTRUCTIONS**

102. I understand that CrowdStrike purports to rely on the plain and ordinary meaning of all claim terms. Pet. at 5 and Bambos at ¶36. For the below key terms, it is helpful, in my opinion, to identify what that plain and ordinary meaning is.

**A. “*social connections*”**

103. In the context of the ’897 patent, a person of ordinary skill in the art would understand “social connections” to refer to “the in-degree and out-degree a given user has to other active users in the social network.”

104. The '897 patent analyzes much data extracted from various darkweb and deepweb sources (collectively “DW”). This data is used to identify “a set of features computed from the social network structure of users posting in a forum.” '897 patent at 11:18-22.

105. These DW Social Network Features “contain[] measures computed from the social connections of users posting hacking-related content. The basic social network features (e.g., in-degree, out-degree) indicates how active a user is in the social graph. More advanced features measure the centrality of users in the social graph. Highly central users are more important; thus the vulnerability mentions should take more consideration. The features were computed for the set of users who explicitly mentioned one or more CVE-IDs in their posts.” '897 patent at 11:29–38. The '897 patent provides great detail about its use of these social network measures. '897 patent at 21:6–47.

106. This social feature information, including “social connection” metrics, are used, for example, to determine “Importance Measures.” '897 patent at 21:31–46. These influence metrics aim to incorporate the “centrality of users” in the relevant social metric to assign weight to communication to or from that user. For example, the patent discloses “Eigenvector centrality” and “Pagerank centrality” to evaluate the importance of the retrieved data.

107. The '897 patent defines in-degree and out-degree as the number of edges pointing to the user (in-degree) and the number of edges originated from the user (out-degree). '897 patent at 21:12–14.

108. Thus, in the context of the '897 patent, a person of ordinary skill in the art would understand “social connections” to refer to the in-degree and out-degree a given user has to other active users in the social network.

***B. “disclosure date”***

109. The '897 patent identifies the disclosure date as the day the National Institute of Standards and Technology (NIST) officially releases the vulnerability number. '897 patent at 12:35–39. “FIG. 3 shows that more than 90% of cases, vulnerabilities are disclosed by NIST before any real-world attack is detected. In the other few cases, the attacks were launched in the wild before NIST published the vulnerabilities.”)

***C. “exploitation date”***

110. The '897 patent defines the “exploitation date” of a vulnerability as the date the exploit was first detected in the wild. '897 patent at 9:42–43.

***D. “likelihood of exploitation”***

111. The '897 patent uses the “likelihood of exploitation” as spectrum, not merely as a binary choice of “exploited” or “not exploited.” For example, the patent discloses that once vulnerabilities are disclosed publicly, “the likelihood of those vulnerabilities being exploited increases.” '897 patent at 1:36–38. A person of

ordinary skill in the art would understand that an “increase” in likelihood of exploitation means something akin in the increase from “low likelihood” to “high likelihood.”

112. This is confirmed in the specification in multiple instances. For example, the prior disclosure teaches that disclosure of a vulnerability on NVD increases the likelihood of exploitation. ’897 patent, table 6. The ’897 patent further teaches that vulnerability mentions on EDB, ZDI, and DW further “increase in the vulnerability exploitation likelihood over vulnerabilities only disclosed on NVD.” ’897 patent at 4:1–4. Thus, the ’897 patent uses a likelihood spectrum, not a binary choice. The ’897 patent even quantifies the increase. “[Our results] demonstrate the likelihood of being exploited given vulnerability mention on EDB is 9%, ZDI is 12%, and DW is 14% as compared to NVD, which is 2.4%.” ’897 patent at 4:5–7.

113. The ’897 patent further discloses “ranking vulnerabilities,” which would be impossible in a binary classification model. ’897 patent at 7:15–17; *see also* ’897 patent at 11:1–3 “The presence of proof-of-concept exploits in EDB increases the likelihood of a vulnerability being exploited.”); *id.* at 11:6–7 (“Vulnerability mention on ZDI also increases the likelihood of it being exploited”); *id.* at 11:11–12 (“Vulnerability mention on DW also increases the likelihood of it being exploited.”).

114. Therefore, in my opinion, a person of ordinary skill in the art would understand the term “likelihood of exploitation” to require multiple classifications

of different likelihoods of exploitation beyond merely the binary choice of “exploited” or “not exploited.”

## VIII. VALIDITY ANALYSIS

115. I understand that CrowdStrike argues the following obviousness grounds:

| Ground | Allegation of Invalidity                                                     | Claims                       |
|--------|------------------------------------------------------------------------------|------------------------------|
| 1      | Obviousness in light of <i>Sabottke</i>                                      | 1, 3, 5, 9–10, 14–15, 18, 20 |
| 2      | Obviousness in light of <i>Smyth</i> and <i>Phillips</i>                     | 1–3, 5, 10–11, 13–15, 17–20  |
| 3      | Obviousness in light of <i>Smyth</i> , <i>Phillips</i> , and <i>Nunes</i>    | 4 and 16                     |
| 4      | Obviousness in light of <i>Smyth</i> , <i>Phillips</i> , and <i>Kuperman</i> | 6                            |
| 5      | Obviousness in light of <i>Smyth</i> , <i>Phillips</i> , and <i>Marin</i>    | 7 and 13                     |
| 6      | Obviousness in light of <i>Smyth</i> , <i>Phillips</i> , and <i>Bozorgi</i>  | 8 and 12                     |

116. Below is my analysis of the asserted prior art references and my opinion as to what they do and do not disclose.

### A. *Ground 1: Obviousness in Light of Sabottke*

117. Petitioner asserts that claims 1, 3, 5, 9-10, 14-15, 18, and 20 are obvious in light of *Sabottke*.

118. The '897 patent cites and distinguishes *Sabottke*. EX1002 at 173 (IDS listing *Sabottke*); EX1001 at 15:35-46 (“The present model was compared to recent works on using vulnerabilities mentioned on Twitter to predict if such vulnerabilities would be exploited or not.”) The '897 patent specifically notes that “there is no existing work on predicting exploits in the wild using DW [DarkWeb] data, and (2)

it is compared with all major approaches, and using feeds from social media currently is the best one.” *Id.*

## 1. Claim 1

**Claim 1.** A method for assessing a likelihood of exploitation of software vulnerabilities, comprising:

*a) Claim 1.b.: accessing a plurality of datasets associated with a predetermined set of data sources, the plurality of datasets including training data comprising hacker communications;*

119. Petitioner identifies “Twitter feeds and the NVD and OSVDB databases” as the “plurality of datasets from predetermined data sources.” Petition at 18, 19.

120. Of the three datasets petitioner identifies Twitter, NVD, and OSVDB, Twitter is the only one that may comprise hacker communications. Neither NVD (National Vulnerability Database) nor OSVDB (Open Sourced Vulnerability Database) comprise hacker communications. Both are vulnerability databases that publish comprehensive lists of vulnerabilities (NVD) and additional attributes.

121. *Sabottke* confirms that it only derives vulnerability information from NVD and OSVDB. *Sabottke* at 3, fn. 7 (referencing NVD (available at <https://nvd.nist.gov/>) and OSVDB (available at <http://oss-security.openwall.org/wiki/mailling-lists/oss-security>)); *see also Sabottke* at 5 (“To obtain context about the vulnerabilities discussed on Twitter, we query the National Vulnerability Database (NVD) [7] for the CVSS scores, the products affected and

additional references about these vulnerabilities. Additionally, we crawl the Open Sourced Vulnerability Database (OSVDB) [9] for a few additional attributes, including the disclosure dates and categories of the vulnerabilities in our study.”).

122. This is also confirmed by the ’897 patent, which discloses that NVD comprises vulnerability information. ’897 patent at 1:60–64 (“NIST provides the National Vulnerability Database (NVD) which comprises of a comprehensive list of vulnerabilities disclosed”).

123. *Sabottke* notes that the advantage of using the Twitter dataset is that it contains information, i.e., hacker communications, that is “available publicly, but is *not included in existing vulnerability databases* such as the National Vulnerability Database (NVD) [7] or the Open Sourced Vulnerability Database (OSVDB).” *Sabottke* at 3.

124. Instead, *Sabottke* confirms that it uses NVD and OSVDB only to gather information about the vulnerabilities discussed in the Twitter dataset.

- a. “To obtain context about the vulnerabilities discussed on Twitter, we query the National Vulnerability Database (NVD) [7] for the CVSS scores, the products affected and additional references about these vulnerabilities. Additionally, we crawl the Open Sourced Vulnerability Database (OSVDB) [9] for a few additional attributes, including the

disclosure dates and categories of the vulnerabilities in our study.”

*Sabottke* at 5.

125. *Sabottke* thus gathers “hacker communications” exclusively from a single dataset gathered from Twitter, and uses these additional vulnerability data sources only to “obtain context about the vulnerabilities discussed on Twitter,” including by querying “the National Vulnerability Database (NVD)” for CVSS scores, products affected, and additional references about these vulnerabilities, and crawling the “Open Sourced Vulnerability Database (OSVDB)” for “disclosure dates and categories of vulnerabilities[.]”] *Id.* Again, the NVD and OSVDB datasets do not comprise hacker communications.

126. In contrast, the ’897 patent gathers “hacker communications” from a much larger and diverse dataset. For example, the ’897 patent discloses “a database of 46 darkweb and deepweb forums.” ’897 patent at 9:50–53. This comes with distinct advantages. ’897 patent at 12:5–6 (“Higher variability between [hacker communication] data sources 110 implies higher information gain and less redundancy.”).

127. Therefore, in my opinion, *Sabottke* does not disclose “a plurality of datasets . . . including training data comprising hacker communications.” *Sabottke* discloses only a single dataset gathered from Twitter that comprises hacker communications. That is precisely the reason why, as the ’897 patent notes, a

comparison of the invention to *Sabottke* using only “vulnerabilities mentioned on Twitter to predict if such vulnerabilities would be exploited or not” demonstrated that “using feeds from social media [i.e., data gathered from Dark Web and Deep Web] currently is the best one.” ’897 patent at 15:42–46. This is further confirmed by another paper, concluding that “models of software vulnerability exploitation based on NVD database entries and social media [Twitter] alone are unlikely to have enough predictive power to be useful in practice.” Benjamin L. Bullough, Anna K. Yanchenko, Christopher L. Smith, and Joseph R. Zipkin, *Predicting Exploitation of Disclosed Software Vulnerabilities Using Open-Source Data*, Proceedings of the 3rd ACM on International Workshop on Security And Privacy Analytics (2017) 45–53 (EX2002, “Zipkin”) at 8–9.

***b) Claim 1.c.: accessing features from the plurality of datasets that include measures computed from social connections of users posting hacking-related content***

128. As explained above, “social connections” refers to the in-degree and out-degree a given user has other active users in the social network. See sec. VII.A. above (construction of “social connections”).

129. Petitioner does not identify any element of *Sabottke* that discloses such “social connections.” In my opinion, *Sabottke* does not disclose any such analysis of “social connections.” This disclosure would require *Sabottke* to analyze the structure of the social network of the authors of the Twitter data it utilizes. This

would have to include, at a minimum, the number of connections (Twitter followers) a given Twitter author has, and the number of others the Twitter user is following. Additional information could include whether the subject tweets considered in *Sabottke* were either responded to by others, or were a response to others. *Sabottke* has no such disclosure, and does not consider the “social connection” of the authors of the Twitter data it considers at all.

130. In my opinion, it would not have been obvious to include such a sophisticated analysis of Twitter author’s social networks at the relevant time of the invention, in part because it requires significantly deeper data collection and analysis than what is proposed by *Sabottke*.

131. Petitioner argues that *Sabottke*’s “Twitter Traffic” features include “measures computed from social connections of users posting hacking-related content. Namely, at least the “# users with minimum T followers/friends,” “# retweets/replies,” and “# tweets favorited” features represent measures computed from the poster’s social connections.” Pet. at 21 (citing *Bambos* Dec., ¶63). None of these disclosures identify or disclose “social connections.” Hence, I disagree.

132. The “# users with minimum T followers/friends” refers to a threshold for user data inclusion. Data from users with no followers/friends is excluded to combat attacks against its exploit detectors. *Sabottke* at 13 (“Cheap Twitter accounts purchased in bulk have low friend counts and low follower counts.”) Furthermore,

the asymmetry of Twitter followers and friends would skew the data, rendering it meaningless: a follower is a one-directional relationship, while a friend is bi-directional; this means that their combination definitely cannot be used to correctly determine at least a Twitter user's social connection in-degree. In turn, this will skew any data derived from such incomplete information. Moreover, neither the "# retweets/replies" nor the "# tweets favorited" identify the Twitter users' "social connections." A user may have many retweets/replies and/or many favorited tweets, but few social connections, and vice versa.

133. It is therefore my opinion that *Sabottke* does not disclose this claim element.

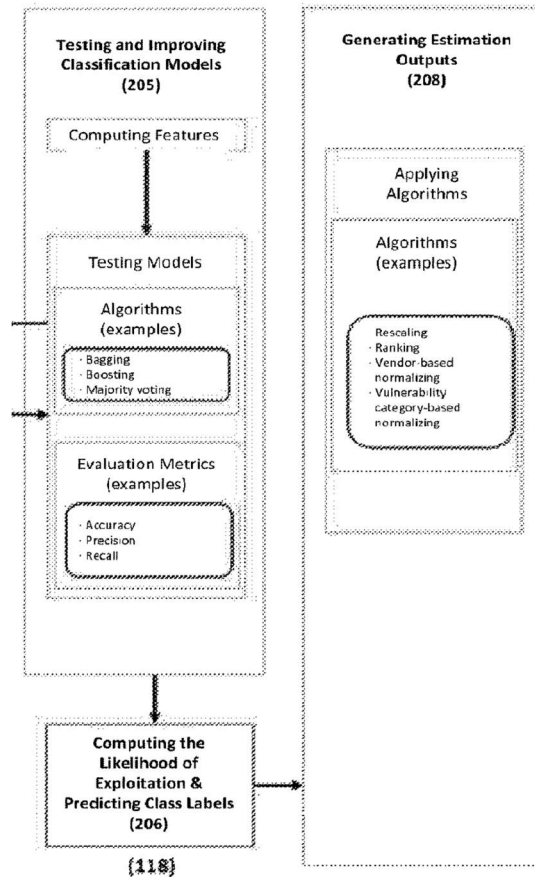
***c) Claim 1.d.: applying learning algorithms to the training data to generate classification models that are configured to predict class labels defining a likelihood of exploitation of respective software vulnerabilities;***

134. Petitioner notes that *Sabottke* "aim[s] to train a classifier that produces binary predictions: each vulnerability is classified as either exploited or not exploited." Pet. at 22. *Sabottke*'s binary classification is insufficient to disclose the claimed "likelihood of exploitation," which enables, for example, the ability to indicate a priority ranking of vulnerabilities likely to be exploited. *See, e.g., '897 patent at 7:15–17.*

135. Petitioner relies on "a class label (e.g., "exploitability") defining exactly two values (e.g., "exploited," "not exploited"). This binary choice presents only

information regarding whether a given vulnerability is expected to be exploited or not. It does not “a likelihood of exploitation,” which instead would require a classification such as “low”, “medium”, or “high” to indicate a probability. The purpose of the invention, after all, is “to prioritize patching.” ’897 patent at abstract and 1:24–28. Classifying a number of vulnerabilities only with the class label of “exploited” does not achieve this goal, since it does not help determine which vulnerability ought to be patched first.

136. This is further corroborated by the embodiment of the taught prediction model in Figure 13 of ’897. Likelihood of Exploitation is computed in block 206, which is fed to Generating Estimation Outputs 208, as shown in the excerpt below. Providing a mere binary distinction between “exploited” and “not exploited” would essentially feed data to the algorithms that capture the likelihood as 0 (i.e., 0% or “will not happen”) or 1 (i.e., 100% or “guaranteed to occur”). This discrete nature of the inputs would directly exclude the embodiments discussed in the ’897 patent’s specification, e.g., increasing the likelihood of exploitation based on the presence of proof-of-concept exploits in EDB or vulnerability mention on ZDI and DW. ’897 patent at 11:1–14. *Sabottke* does not allow “increasing the likelihood” as taught in the ’897 patent, but rather either strictly excludes the vulnerability’s possibility, or guarantee its exploitation.



'897 patent, fig. 13 (excerpt).

137. For these reasons, it is my opinion that *Sabottke's* discreet choices of “exploited” and “not exploited” do not express the claimed “likelihood of exploitation.”

***d) Claim 1.f.: computing, by applying the one or more features to the classification model, a class label defining one or more values defining a likelihood of exploitation associated with the software vulnerability,***

138. As stated regarding claim element 1.d, Petitioner relies on “a class label (e.g., “exploitability”) defining one or more values (e.g., “exploited,” “not exploited”). See element 1.d. This is a binary choice: guaranteed not to happen vs.

guaranteed to happen. Pet. at 24 (“A vulnerability receives an exploitability class label that defines a binary value—“exploited, not exploited.”) As explained above, that does not disclose the claimed “likelihood of exploitation.” All vulnerabilities with the class label “exploited” are treated equally by *Sabottke*, and no likelihood value is assigned.

139. This is corroborated by Figure 13 of ’897. The Likelihood of Exploitation is computed in block 206, which is fed to Generating Estimation Outputs 208, as shown in the excerpt below. Providing a mere binary distinction between “exploited” and “not exploited” would essentially feed data to the algorithms that capture the likelihood as 0 (i.e., 0% or “will not happen”) or 1 (i.e., 100% or “guaranteed to occur”). This discrete nature of the inputs would directly exclude the embodiments discussed in the ’897 patent’s specification, e.g., increasing the likelihood of exploitation based on the presence of proof-of-concept exploits in EDB or vulnerability mention on ZDI and DW. ’897 patent at 11:1–14. *Smyth* does not allow “increasing the likelihood” as taught in ’897, but rather either strictly excludes the vulnerability’s possibility, or guarantee its exploitation.

140. The ’897 patent confirms in many instances that “likelihood of exploitation” is a gradient, as explained above in the construction of this claim term. *See, e.g.*, ’897 patent, table 6 (mention on NVD increases the likelihood of exploitation); *id.* at 4:1–4 (vulnerability mentions on EDB, ZDI, and DW further

“increase in the vulnerability exploitation likelihood over vulnerabilities only disclosed on NVD.”) Thus, the ’897 patent uses a likelihood spectrum, not a binary choice. The ’897 patent even quantifies the stepwise increase. ’897 patent at 4:5–7.

141. The ’897 patent further discloses “ranking vulnerabilities,” which would be impossible in a binary classification model. ’897 patent at 7:15–17.

***e) Claim 1.f.1.: wherein the likelihood of exploitation predicts an actual exploitation of the respective software vulnerabilities before disclosure based on the hacker communications from the training data.***

142. *Sabottke* explicitly states that “we do not consider the detection of zero-day attacks [32], which exploit vulnerabilities before their public disclosure; instead, we focus on detecting the use of exploits against known vulnerabilities.” *Sabottke* at 3. Thus, *Sabottke* does not disclose predicting “an actual exploitation of the respective software vulnerabilities before disclosure” (emphasis added).

143. Petitioner cites several aspects of *Sabottke* to counter this fact. First, Petitioner argues that “*Sabottke* expressly teaches methods to ‘characterize the landscape of threats related to information leaks about vulnerabilities before their public disclosure’ to a public vulnerability database.” Pet. at 25–26 (citing *Sabottke* at 3). This statement does not support that *Sabottke* “predicts an actual exploitation of the respective software vulnerabilities before disclosure.” Instead, it refers to “characterize[ing] the landscape of threats” in order to “identify features that can be

extracted automatically from the Twitter discourse to detect exploits.” *Sabottke’s* reference to identifying features “before disclosure” does not support that it predicts actual exploitation “before disclosure.”

144. Second, Petitioner cites *Sabottke’s* statement that it identified “47 vulnerabilities that were mentioned in the Twitter stream before their public disclosure.” Pet. at 26 (citing *Sabottke* at 9). Again, this statement does not support that *Sabottke* “predicts an actual exploitation of the respective software vulnerabilities before disclosure.” Instead, it merely refers to Twitter references to a vulnerability before disclosure thereof. *Sabottke* then “investigate these cases manually to determine the sources of this information.” *Sabottke* at 9. Nothing in this section supports exploit prediction before disclosure, let alone doing so algorithmically.

145. Third, Petitioner cites *Sabottke’s* section “5.1 Early Detection of Exploits.” Pet. at 13. In that section, *Sabottke* acknowledges at the outset that “[w]hile the first tweets about a vulnerability precede the creation of IPS or AV signatures by a median time of 10 days, these first tweets are typically not informative enough to determine that the vulnerability is likely exploited.” *Sabottke* at 12 (emphasis added). That statement confirms that *Sabottke* does not teach pre-disclosure exploit prediction.

146. Further, *Sabottke* then discloses that—because “these first tweets are typically not informative enough”—they created “10 stratified random samples, each with 50% coverage of “exploited” and “not exploited” vulnerabilities” in order to “train a separate linear SVM classifier.” *Sabottke* at 13. This disclosure refers to creating test-data sets based on tweets before and after disclosure of the vulnerability, in order to train classifiers that might detect a vulnerability as close to the first Twitter mention as possible. *Id.* In other words, *Sabottke* utilizes post-exploitation training data. This is specifically pointed out in other papers. *See, e.g., Zipkin* at 3 (“*Edkrantz* [5] and *Sabottke* [13] each randomly split training and test data, allowing the temporal intermixing of “future” and “past” data in the training and test sets.”); *see also* Haipeng Chen, et al., *Using Twitter to Predict When Vulnerabilities will be Exploited*, KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Pages 3143 – 3152 (EX2003, “Chen”) at 3144 (“Because 10-fold cross validation splits data randomly into 10-folds, the 9 folds used for training will almost certainly contain data from the future, and then use those to ‘predict’ some exploits used in the past.”)

147. Specifically, for this simulation, *Sabottke* chose “an aggregated prediction threshold of 0.95, which achieves 45% precision.” *Sabottke* at 13. “Our Twitter-based classifier provides an “exploited” output 3 hours after the first tweet appears related to this CVE on April 7, 2014.” *Id.* *Sabottke* can arbitrarily modify the time

of detection after the first tweet by lowering its precision threshold. *Id.* (“By accepting lower levels of precision, our Twitter-based classifiers can achieve even faster exploit detection. For example, with classifier precision set to approximately 25%, Heartbleed can be detected as an exploit within 10 minutes of its first appearance on Twitter.”) This flexibility is possible only because *Sabottke* does not train its classifiers on only pre-disclosure data. Instead, the random data set *Sabottke* uses to train its classifiers include tweets both before and after disclosure. This in turn means that what *Sabottke* is disclosing here is not pre-disclosure vulnerability detection based on pre-disclosure tweets. Instead, *Sabottke* expressly discloses creating a classifier based on both pre- and post-disclosure tweets, and then testing to see if that classifier *could have* detected the vulnerability *if it had been available before disclosure*.

148. The ’897 patent specifically discourages this practice: “exploit prediction is a time dependent prediction problem. Hence splitting the data randomly violates the temporal aspect of the data as events that happen in the future will now be used to predict events that happen in the past, which violates the sequential nature of vulnerability disclosure events. This leads to temporal intermixing of the data.” ’897 patent at 5:15–25.; *see also id.* at 16: 21–26 (“Avoiding temporal intermixing. It should be understood that the temporal sequencing of exploit warning could lead to future events being used to predict past ones. This could lead to inaccurate results,

especially in the case where there is strong temporal dependencies in the data-which argues is the case for social media.”).

149. Since *Sabottke*’s classifiers are not available *before* disclosure (they are created only after disclosure, based in part of post-disclosure tweets), *Sabottke* does not disclose pre-disclosure prediction.

## **2. Claim 3**

**Claim 3.** The method of claim 1, wherein the plurality of datasets include vulnerability data for vulnerabilities that are publicly disclosed and obtaining exploits data for exploits that were used in real world attacks.

150. Claim 3 depends on Claim 1 and is not disclosed by *Sabottke* for the same reasons as stated in regard to Claim 1.

## **3. Claim 5**

**Claim 5.** The method of claim 1, wherein certain features correspond to a known vulnerability obtained from the plurality of datasets.

151. Claim 5 depends on Claim 1 and is not disclosed by *Sabottke* for the same reasons as stated in regard to Claim 1.

## **4. Claim 9**

**Claim 9.** The method of claim 1, further comprising computing mutual information from the plurality of datasets informative as to what information a given feature provides about another feature.

152. Claim 9 depends on Claim 1 and is not disclosed by *Sabottke* for the same reasons as stated in regard to Claim 1.

## 5. Claim 10

**Claim 10.** The method of claim 1, further comprising:

detecting, from the plurality of datasets, vulnerabilities that appear before an associated exploitation date.

153. The '897 patent defines the “exploitation date” of a vulnerability as the date the exploit was first detected in the wild. '897 patent at 9:42–43. *Sabottke* is incapable of meeting and does not disclose this requirement.

154. *Sabottke* explicitly states that “we do not consider the detection of zero-day attacks [32], which exploit vulnerabilities before their public disclosure; instead, we focus on detecting the use of exploits against known vulnerabilities.” *Sabottke* at 3. Thus, *Sabottke* does not disclose detecting “vulnerabilities that appear before an associated exploitation date.”

## 6. Claim 14

**Claim 14:** A computing device, configured via machine learning to apply a learned function to data associated with a software vulnerability to estimate a likelihood of exploitation of the software vulnerability, the learned function associated with a prediction model derived from at least one machine learning algorithm and a plurality of datasets associated with software vulnerabilities, the plurality of datasets including information associated with discussion of vulnerabilities by users posting hacking-related content, wherein the likelihood of exploitation predicts an actual exploitation of the software vulnerability before disclosure based on hacker communications from the plurality of datasets.

155. Petitioner and Dr. Bambos only cursorily address claim 14 without addressing all elements of claim 14. Pet. at 31, 32; Bambos at ¶¶ 79, 80.

156. Petitioner cites its claim 1 analysis generally for claim 14 applying “a learned function to data associated with a software vulnerability to estimate a likelihood of exploitation of the software vulnerability.” Pet. at 31. I therefore incorporate my corresponding response to claim 1. *See* section VIII.A.1.

157. Petitioner incorporates its analysis of claim elements 1.f and 1.f.1. regarding claim elements “the plurality of datasets including information associated with discussion of vulnerabilities by users posting hacking-related content” and “wherein the likelihood of exploitation predicts an actual exploitation of the software vulnerability before disclosure based on hacker communications from the plurality of datasets.” Pet. at 31–32. I therefore incorporate my corresponding response thereto.

158. For the reasons stated regarding Claim 1, *Sabottke* does not disclose a “plurality of datasets including information associated with discussion of vulnerabilities by users posting hacking-related content” (*see* claim element 1.b); “the likelihood of exploitation” (*see* claim element 1.f); and predicting “an actual exploitation of the software vulnerability before disclosure” (*see* claim element 1.f.1).

159. Because the indicated claim limitations are not disclosed by *Sabottke*, it is my opinion that *Sabottke* does not teach or suggest claim 14 to a person of ordinary skill in the art at the time of the invention.

## 7. Claim 15

**Claim 15.** The computing device of claim 14, wherein the prediction model is at least one classification model that outputs from features of the software vulnerability a score indicative of the likelihood of exploitation.

160. Petitioner cites 1.f regarding this claim. Pet. at 32. I therefore incorporate my response thereto. Specifically, Sabottke discloses only a binary categorization of “exploited” and “not exploited,” and therefore does not disclose the gradient “likelihood of exploitation” claimed in the ’897 patent.

## 8. Claim 18

**Claim 18:** A method of prioritizing vulnerabilities using cyber threat intelligence, comprising:

*a) Claim 18.c.: applying learning algorithms to the training data to generate a predictive model configured to predict a likelihood of exploitation of respective software vulnerabilities;*

161. Petitioner addresses this claim element only by reference to claim element 1.d. I therefore incorporate my analysis of that claim element. For the reasons stated there, *Sabottke* does not disclose this claim element, and specifically, the claimed “likelihood of exploitation.”

***b) Claim 18.e.: computing, by applying the one or more features to the predictive model, one or more values defining a likelihood of exploitation associated with the software vulnerability, wherein the likelihood of exploitation predicts an actual exploitation of the respective software vulnerabilities before disclosure based on the hacker communications from the training data.***

162. Petitioner addresses this claim element only by reference to claim element 1.f and 1.f.1. I therefore incorporate my analysis of those claim elements. For the reasons stated there, *Sabottke* does not disclose this claim element, and specifically, the claimed “likelihood of exploitation” and predicting “an actual exploitation . . . before disclosure.”

## **9. Claim 20**

**Claim 20.** The method of claim 18, wherein the plurality of datasets include vulnerability data for vulnerabilities that are publicly disclosed and obtaining exploits data for exploits that were used in real world attacks.

163. Claim 20 depends on Claim 18 and is not disclosed by *Sabottke* for the same reasons as stated in regard to Claim 18.

### ***B. Ground 2: Obviousness in Light of Smyth and Phillips***

164. Petitioner asserts that claims 1-3, 5, 10-11, 13-15, 17-20 are obvious in light of *Smyth* (EX1005) and *Phillips* (EX1006).

165. The ’897 patent cites *Phillips*. EX1002 at 173 (IDS listing *Phillips*); *see also* EX1001 at 2.

## 1. Claim 1

**Claim 1.** A method for assessing a likelihood of exploitation of software vulnerabilities, comprising:

*a) Claim 1.b.: accessing a plurality of datasets associated with a predetermined set of data sources, the plurality of datasets including training data comprising hacker communications;*

166. Petitioner relies on *Phillips* to supplement the disclosure of *Smyth* with the claimed “hacker communications.” Pet. at 38. I disagree for multiple reasons.

167. First, *Phillips* is not analogous art. Unlike the ’897 patent, *Phillips* explores “various Social Network Analysis (SNA) techniques in order to identify a range of potentially ‘important’ members of Islamic Networks within Dark Web Forums.” *Phillips* at 97. *Phillips*’ only aim is to identify individuals. It disregards the content of the communications and relies exclusively on communication metadata. *Id.* *Phillips* is not in the computer network vulnerability area at all.

168. Second, despite being familiar with “state of-the-art Computational Linguistic techniques” for analyzing communication content, *Phillips* intentionally did not do so. *Phillips* at 97 (“Whilst existing work attempts to identify these structures through state of-the-art Computational Linguistic techniques, our work relies on the communication metadata alone.”). That is because use of communication content—i.e., “hacker communications”—is a complex task. *Phillips* specifically notes the difficulty of scaling a system that relies on communication content. *Id.* (“Whilst existing research has been undertaken into

creating a social network topology for Dark web Forums [11], these existing research focus on small communities and networks and rely on sophisticated sentiment analysis techniques, which are difficult to scale.”). *Phillips* at 97. Thus, *Phillips* does not collect or consider “hacker communications.” To the contrary, *Phillips* teaches away from relying on communication content when scaling the analysis to larger datasets, due to the complexity of the analysis.

169. For these reasons, it is my opinion that the combination of *Smyth* and *Phillips* does not teach or suggest this claim element to a person of ordinary skill in the art.

170. Further, for the reasons detailed in the “Motivation to Combine” section below, a person of ordinary skill in the art would not have looked to *Phillips* to supplement the disclosure of *Smyth* to arrive at this claim element.

***b) Claim 1.c.: accessing features from the plurality of datasets that include measures computed from social connections of users posting hacking-related content***

171. Petitioner admits that “*Smyth* does not expressly discuss accessing features that include “measures computed from social connections of users posting hacking-related content.” Pet. at 41. Neither does *Phillips*.

172. *Phillips* discloses using “Social Network Analysis (SNA) techniques” to analyze dark web forums. However, “social connections” of users posting hacking-

related content is not one of the metrics *Phillips* considered. Petitioner’s citation to Table 1 of *Phillips* is illustrative:

TABLE I. A TABLE OUTLINING THE METRICS USED TO EVALUATE OUR SOCIAL NETWORK

| Attribute Name                     | Description                                                                                                                                              |
|------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sent Messages (SM)                 | The number of emails sent by an employee.                                                                                                                |
| Received Messages (RM)             | The number of emails received by an employee.                                                                                                            |
| Degree Centrality (DCS)            | The number of distinct employees within the network that an employee has sent emails to.                                                                 |
| Betweenness Centrality Score (BCS) | The betweenness centrality measure for an employee.[20]                                                                                                  |
| Pagerank Score (PRS)               | The PageRank score an employee. [21]                                                                                                                     |
| Markov Ranking (MR)                | The markov ranking of an employee. [22]                                                                                                                  |
| HITS Authority Score (HAS)         | The authority score for an employee (if several users with high hub weights send an email t the user then they will have a higher authority score). [23] |
| HITS Hub Score (HHS)               | The hub score for an employee (if the user sends emails to users with high authority scores then they will have a higher hub score). [23]                |
| Clique Score (CS)                  | The number of cliques (maximal subgraphs) an employee is in using the Bron and Kerbosch algorithm.[24]                                                   |
| Weighted Clique Score (WCS)        | The weighted clique score for each user, weighted by the number of users within each clique.                                                             |
| Average Distance Score (ADS)       | The average distance between the user and all other users in the graph.                                                                                  |
| Clustering Coefficient (CC)        | The extent to which vertices in a graph tend to cluster together. [25]                                                                                   |

Pet. at 41 (citing *Phillips*, Table 1).

173. The only metrics considered (highlighted by Petitioner), are number of emails sent, number of emails received, number of distinct employees that an employee has sent emails to, the betweenness centrality measure for an employee, and the PageRank score of an employee. *Id.* None of these metrics are “social connections,” i.e., the in-degree and out-degree a given user has to other active users in the social network. For example, sending a large number of emails to very few of your connections and/or receiving many emails from a small number of connections would significantly skew *Phillips*’ results.

174. Ultimately, Petitioner argues that “*Phillips* teaches collecting dark web forum posts, extracting metadata from these posts, building a social network, and then analyzing that social network with SNA techniques and machine learning.” Pet.

at 42. Even assuming this is true, it does not disclose “measures computed from social connections of users posting hacking-related content.”

***c) Claim 1.d.: applying learning algorithms to the training data to generate classification models that are configured to predict class labels defining a likelihood of exploitation of respective software vulnerabilities;***

175. Petitioner argues that *Smyth*’s “prediction ensemble 155 comprises trained algorithms 531–534 and therefore comprises ‘classification models’ as claimed. Pet. at 46 (citing Bambos at ¶113). However, Dr. Bambos only offers the same conclusory statement: “Prediction ensemble 155 comprises trained algorithms 531–534 and therefore comprises ‘classification models’ as claimed.” Bambos at ¶ 113. Petitioner then relies on the ’897 patent itself to provide “provides exemplary classification models that predict ‘class labels.’” Pet. at 46. Notably, *Smyth* does not disclose any classification models, nor any class labels, and Petitioner does not identify any element in *Smyth* to meet these elements beyond training the AI engine generally. I disagree with this argument because neither classification models nor class labels are necessarily inherent in training an AI engine.

176. The fact that *Smyth* ultimately aims to predict the likelihood of exploitation of system vulnerabilities (*see Smyth* Fig. 1) does not disclose how *Smyth* accomplishes this task. Use of classification models and class labels are not necessarily required and can therefore not be inferred.

***d) Claim 1.f.: computing, by applying the one or more features to the classification model, a class label defining one or more values defining a likelihood of exploitation associated with the software vulnerability,***

177. Petitioner argues that “*Smyth*’s prediction ensemble computes the exploitability of a vulnerability. The class label is therefore the variable ‘exploitability’ and that variable is defined by values such as ‘exploited’ or ‘not exploited.’” Pet. at 50 (citing *Bambos* at ¶118). I disagree for multiple reasons.

178. First, as stated regarding element 1.d, *Smyth* does not explicitly disclose either classification models nor class labels, and neither is necessarily inherent in merely training an AI engine.

179. Second, as stated regarding claim element 1.d in ground 1 (regarding *Sabottke*), Petitioner relies on the binary class label “exploitability,” which has only two values, “exploited” and “not exploited.” Pet. at 50. This is a binary choice: guaranteed-not-to-happen vs. guaranteed-to-happen. That binary choice does not disclose the claimed “likelihood of exploitation” of the ’897 patent, because that term requires a gradient. All vulnerabilities with the class label “exploited” are treated equally by Petitioner’s interpretation of *Smyth*, and no likelihood value is assigned.

180. This is corroborated by Figure 13 of ’897. The Likelihood of Exploitation is computed in block 206, which is fed to Generating Estimation Outputs 208, as shown in the excerpt below. Providing a mere binary distinction between

“exploited” and “not exploited” would essentially feed data to the algorithms that capture the likelihood as 0 (i.e., 0% or “will not happen”) or 1 (i.e., 100% or “guaranteed to occur”). This discrete nature of the inputs would directly exclude the embodiments discussed in the ’897 patent’s specification, e.g., increasing the likelihood of exploitation based on the presence of proof-of-concept exploits in EDB or vulnerability mention on ZDI and DW. ’897 patent at 11:1–14. *Smyth* does not allow “increasing the likelihood” as taught in ’897, but rather either strictly excludes the vulnerability’s possibility, or guarantee its exploitation.

181. The ’897 patent confirms in many instances that “likelihood of exploitation” is a gradient. *See, e.g.*, ’897 patent, table 6; *id.* at 4:1–4. Thus, the ’897 patent uses a likelihood spectrum, not a binary choice. The ’897 patent even quantifies the stepwise increase. ’897 patent at 4:5–7.

182. The ’897 patent further discloses “ranking vulnerabilities,” which would be impossible in a binary classification model. ’897 patent at 7:15–17.

***e) Claim 1.f.1.: wherein the likelihood of exploitation predicts an actual exploitation of the respective software vulnerabilities before disclosure based on the hacker communications from the training data.***

183. *Smyth* does not refer to the “disclosure” of vulnerabilities. Therefore, *Smyth* does not identify whether its disclosures refer to vulnerabilities discovered before or after the disclosure date, i.e., the day the National Institute of Standards

and Technology (NIST) officially releases the vulnerability number. '897 patent at 12:35–39.

184. Petitioner relies on *Smyth*'s reference to “a threat that is newly discovered or currently spreading through computer networks.” Pet. at 51 (citing *Smyth* at 11:1-8 and 14:55–57). However, *Smyth* does not reference the disclosure date at all. Petitioner's inference that reference to a “new threat” necessarily implies disclosure before the priority date is unsupported.

185. Petitioner further cites Dr. Bambos in support. Pet. at 51 (citing Bambos at ¶¶119–123). However, Dr. Bambos admits that “some vulnerabilities are disclosed immediately upon discovery.” Bambos at ¶121 (citing *Bozorgi* at 2). Thus, no inference can be drawn that *Smyth*'s reference to “a threat that is newly discovered or currently spreading through computer networks” must necessarily occur before disclosure. Without more information regarding the specific vulnerability, a person of ordinary skill in the art would not know whether “a newly discovered threat” found in online discussions had already been disclosed “immediately upon discovery.”

## **2. Claim 2**

**Claim 2.** The method of claim 1, further comprising generating a plurality of estimation outputs based on the one or more values to derive an overall quantitative score.

186. Claim 2 depends on Claim 1 and is not disclosed by *Smyth* and *Phillips* for the same reasons as stated in regard to Claim 1.

### **3. Claim 3**

**Claim 3.** The method of claim 1, wherein the plurality of datasets include vulnerability data for vulnerabilities that are publicly disclosed and obtaining exploits data for exploits that were used in real world attacks.

187. Claim 3 depends on Claim 1 and is not disclosed by *Smyth* and *Phillips* for the same reasons as stated in regard to Claim 1.

### **4. Claim 5**

**Claim 5.** The method of claim 1, wherein certain features correspond to a known vulnerability obtained from the plurality of datasets.

188. Claim 5 depends on Claim 1 and is not disclosed by *Smyth* and *Phillips* for the same reasons as stated in regard to Claim 1.

### **5. Claim 10**

**Claim 10.** The method of claim 1, further comprising:

detecting, from the plurality of datasets, vulnerabilities that appear before an associated exploitation date.

189. The '897 patent defines the “exploitation date” of a vulnerability as the date the exploit was first detected in the wild. '897 patent at 9:42–43. *Smyth* makes no reference to the exploitation date, and thus does not clarify whether it detects “vulnerabilities that appear before an associated exploitation date,” or how it could make such a detection.

190. Petitioner references only *Smyth*'s reliance on "current vulnerability information 125 contain[ing] vulnerability information about a threat that is newly discovered" and "more current information such as online discussions 240." Pet. at 55. However, *Smyth* does not reference the exploitation date. It is therefore unclear whether the current vulnerability information 125 and online discussions 240 in *Smyth* are indeed before or after an associated exploitation date. It is possible, and indeed likely, that a vulnerability is discovered and discussed online only because of an exploitation of that vulnerability.

191. Therefore, and for the reasons detailed regarding element 1.f.1., in my opinion, the combination of *Smyth* and *Phillips* does not disclose "detecting . . . vulnerabilities that appear before an associated exploitation date."

## **6. Claim 11**

**Claim 11.** The method of claim 1, further comprising:

accessing features from the plurality of datasets that measure a centrality of the users in a social graph.

192. Claim 11 depends on Claim 1 and is not disclosed by *Smyth* and *Phillips* for the same reasons as stated in regard to Claim 1.

## **7. Claim 13**

**Claim 13.** The method of claim 1, further comprising vectorizing text features derived from textual content of the plurality of datasets using a predetermined natural language process (NLP [sic]).

193. Claim 13 depends on Claim 1 and is not disclosed by *Smyth* and *Phillips* for the same reasons as stated in regard to Claim 1.

## 8. Claim 14

**Claim 14:** A computing device, configured via machine learning to apply a learned function to data associated with a software vulnerability to estimate a likelihood of exploitation of the software vulnerability, the learned function associated with a prediction model derived from at least one machine learning algorithm and a plurality of datasets associated with software vulnerabilities, the plurality of datasets including information associated with discussion of vulnerabilities by users posting hacking-related content, wherein the likelihood of exploitation predicts an actual exploitation of the software vulnerability before disclosure based on hacker communications from the plurality of datasets.

194. Petitioner and Dr. Bambos only cursorily address claim 14 without addressing all elements of the claim. Pet. at 58, 59; Bambos at ¶125.

195. Petitioner cites its claim 1 analysis generally for claim 14 applying “a learned function to data associated with a software vulnerability to estimate a likelihood of exploitation of the software vulnerability.” Pet. at 58. I therefore incorporate my corresponding response to claim 1. *See* section VIII.A.1.

196. Petitioner incorporates its analysis of claim elements 1.f and 1.f.1. regarding claim elements “the plurality of datasets including information associated with discussion of vulnerabilities by users posting hacking-related content” and “wherein the likelihood of exploitation predicts an actual exploitation of the software vulnerability before disclosure based on hacker communications from the plurality

of datasets.” Pet. at 58–59. I therefore incorporate my corresponding response thereto.

197. For the reasons stated regarding Claim 1, the combination of *Smyth* and *Phillips* does not disclose a “plurality of datasets including information associated with discussion of vulnerabilities by users posting hacking-related content” (*see* claim element 1.b); “classification models” and “class labels” (*see* claim element 1.d); “the likelihood of exploitation” (*see* claim element 1.f); and predicting “an actual exploitation of the software vulnerability before disclosure” (*see* claim element 1.f.1).

198. Because the indicated claim limitations are not disclosed by the combination of *Smyth and Phillips*, it is my opinion that the combination of *Smyth and Phillips* does not teach or suggest claim 14 to a person of ordinary skill in the art at the time of the invention.

## **9. Claim 15**

**Claim 15.** The computing device of claim 14, wherein the prediction model is at least one classification model that outputs from features of the software vulnerability a score indicative of the likelihood of exploitation.

199. For the reasons stated regarding claim element 1.d, the combination of *Smyth* and *Phillips* does not disclose a classification model.

200. For the reasons stated regarding claim element 1.f, the combination of *Smyth* and *Phillips* does not disclose a likelihood of exploitation.

## 10. Claim 17

**Claim 17.** The computing device of claim 14, wherein the computing device is further configured to apply the learned function to data associated with each of a plurality of new software vulnerabilities and compute a respective likelihood of exploitation for each of the plurality of new software vulnerabilities.

201. For the reasons stated regarding claim element 1.f, the combination of *Smyth* and *Phillips* does not disclose a “likelihood of exploitation.”

## 11. Claim 18

**Claim 18:** A method of prioritizing vulnerabilities using cyber threat intelligence, comprising:

*a) Claim 18.b: accessing a plurality of datasets associated with a predetermined set of data sources, at least a portion of the plurality of datasets defining training data including hacker communications;*

202. Petitioner addresses this claim element only by reference to claim element 1.b. I therefore incorporate my analysis of that claim element. For the reasons stated there, the combination of *Smyth* and *Phillips* does not disclose this claim element, and specifically, the claimed “training data including hacker communications.”

*b) Claim 18.c: applying learning algorithms to the training data to generate a predictive model configured to predict a likelihood of exploitation of respective software vulnerabilities;*

203. Petitioner addresses this claim element only by reference to claim element 1.d. I therefore incorporate my analysis of that claim element. For the reasons stated

there, the combination of *Smyth* and *Phillips* does not disclose this claim element, and specifically, the claimed “predict[ing] a likelihood of exploitation.”

*c) Claim 18.e: computing, by applying the one or more features to the predictive model, one or more values defining a likelihood of exploitation associated with the software vulnerability, wherein the likelihood of exploitation predicts an actual exploitation of the respective software vulnerabilities before disclosure based on the hacker communications from the training data.*

204. Petitioner addresses this claim element only by reference to claim element 1.d. I therefore incorporate my analysis of that claim element. For the reasons stated there, the combination of *Smyth* and *Phillips* does not disclose this claim element, and specifically, the claimed “predict[ing] a likelihood of exploitation” (*see* claim element 1.f), “before disclosure” (*see* claim element 1.f.1), and “based on hacker communications” (*see* claim element 1.b).

## **12. Claim 19**

**Claim 19.** The method of claim 18, further comprising generating a plurality of estimation outputs based on the one or more values to derive an overall quantitative score.

205. Claim 19 depends on Claim 18 and is not disclosed by *Smyth* and *Phillips* for the same reasons as stated in regard to Claim 18.

## **13. Claim 20**

**Claim 20.** The method of claim 18, wherein the plurality of datasets include vulnerability data for vulnerabilities that are publicly disclosed and obtaining exploits data for exploits that were used in real world attacks.

206. Claim 20 depends on Claim 18 and is not disclosed by *Smyth* and *Phillips* for the same reasons as stated in regard to Claim 18.

#### **14. Motivation to Combine**

207. *Smyth* identifies no problem with its system that *Phillips* would solve. Indeed, *Smyth* and *Phillips* are not aimed at the same goal. *Smyth* aims to predict the likelihood of vulnerability exploitation. *Phillips*, on the other hand, aims to identify ‘important’ members of a given discussion forum based on metadata analysis. Thus, there would have been no reason for a POSITA implementing *Smyth* to look to *Phillips*. Even if *Smyth* had sought to analyze “hacker communications”—and there is no disclosure in *Smyth* to support this assumption—*Phillips* would not have provided a solution because *Phillips* teaches away from analyzing the actual content of the communications, and relies exclusively on metadata. Metadata, by definition, does not comprise “hack communications.” Finally, as explained in *Phillips*, adding analysis of the content of the communications (i.e., hacker communications) was not a trivial task, and would have been complex to implement and even harder to scale. *Phillips* at 97 (analysis of communication content relies on “sophisticated sentiment analysis techniques, which are difficult to scale.”)

208. Petitioner cites EX1013 (“*Benjamin*”) for the proposition that “a rich repository of current hacker communications for *Smyth*’s predictive engine to analyze” can be found in dark web forums. Pet. at 39. While true, that does not

provide the means to perform this complex task. To the contrary, despite the intuitive knowledge that dark web hacker forums contain information useful for threat analysis, *Phillips* teaches that it can meet its goal of identifying important group members without the difficult task of communication content analysis—by relying exclusively on metadata. Thus, *Phillips* teaches that incurring the cost of analyzing “hacker communications” is unnecessary because metadata analysis is sufficient to accomplish its goal. Notably, *Benjamin* also does not consider the communication content. (*Benjamin* at 7 (considering only Average-Message-Length, Number-Of-Replies, Number-Of-Threads-Involved, Tenure, Sum-Of-Attachments, and Total-Messages)).

209. Petitioner states that “it would have been well within the skill of a POSITA to modify *Smyth* to examine hacker communications in dark web forums.” Pet at 40. This is conclusory and does not consider the stark difference between communications of security professionals that openly share ideas, and hacker communications that intentionally obfuscate their clandestine communications. *MacDonald* (EX1015) (“Hacker communities are dark networks, so members explicitly protect their anonymity, which serves to counter detection from law enforcement.”) I therefore agree with *Smyth* and *Benjamin* that the analysis of the substance of hacker communications was a complex and non-obvious task at the relevant time.

210. Furthermore, *Phillips* targets communications in Islamic Networks (e.g., motivated by the 911 terrorist attacks). It has nothing to do with cybersecurity and vulnerable software, and is not analogous art.

211. For these reasons, it is my opinion that a person of ordinary skill in the art would not have been motivated to combine *Smyth* with *Phillips*.

**C. Ground 3: Obviousness in light of *Smyth*, *Phillips*, and *Nunes***

212. Petitioner asserts that claims 4 and 16 are obvious in light of *Smyth*, *Phillips*, and *Nunes*.

**1. Claim 4**

**Claim 4.** The method of claim 3, further comprising:

aligning the exploits data with the vulnerability data; and

cleaning the exploits data of noise and predetermined portions of the exploits data that is irrelevant to associated software vulnerabilities.

213. Petitioner admits that the combination of *Smyth* and *Phillips* “does not expressly describe further processing that exploits data to remove noise and irrelevant data.” Pet. at 65. Petitioner then cites *Nunes* as teaching “further processing that exploits data to remove noise and irrelevant data.” *Id.*

214. In my opinion, a person of ordinary skill in the art would not have been motivated to further combine the combination *Smyth* and *Phillips* with *Nunes*. As explained in regard to claim 1, the combination of *Smyth* and *Phillips* ignores the communications and uses only metadata. See claim ground 2, claim element 1.b.

215. Specifically, Petitioner recognizes that “*Smyth* does not expressly discuss accessing ‘hacker communications’ in those forums.” Pet. at 38. Similarly, *Phillips* only aim is to identify “important individuals” using “communication metadata alone.” *Phillips* at 97. *Phillips* thus disregards the content of the communications. Thus, neither *Smyth* nor *Phillips* would have a need for *Nunes*’ disclosure of further processing the communication content of online discussions to separate relevant data from noise.

216. To the contrary, *Phillips* teaches away from utilizing *Nunes* because *Phillips* intentionally avoided processing communication content. *Phillips* at 97 (“Whilst existing research has been undertaken into creating a social network topology for Dark web Forums [11], these existing research focus on small communities and networks and rely on sophisticated sentiment analysis techniques, which are difficult to scale.”).

217. For these reasons, it is my opinion that the combination of *Smyth*, *Phillips*, and *Nunes* does not teach or suggest claim 4 to a person of ordinary skill in the art.

## **2. Claim 16**

**Claim 16.** The computing device of claim 14, wherein the predictive model includes a Random Forest (RF) method including multiple decision tree predictors applied in combination to classify the software vulnerability which is used to estimate the likelihood of exploitation.

218. Petitioner cites to its analysis of ground 2, claim 14, and accordingly, I incorporate my response thereto here. However, Petitioner notes that “*Smyth* does

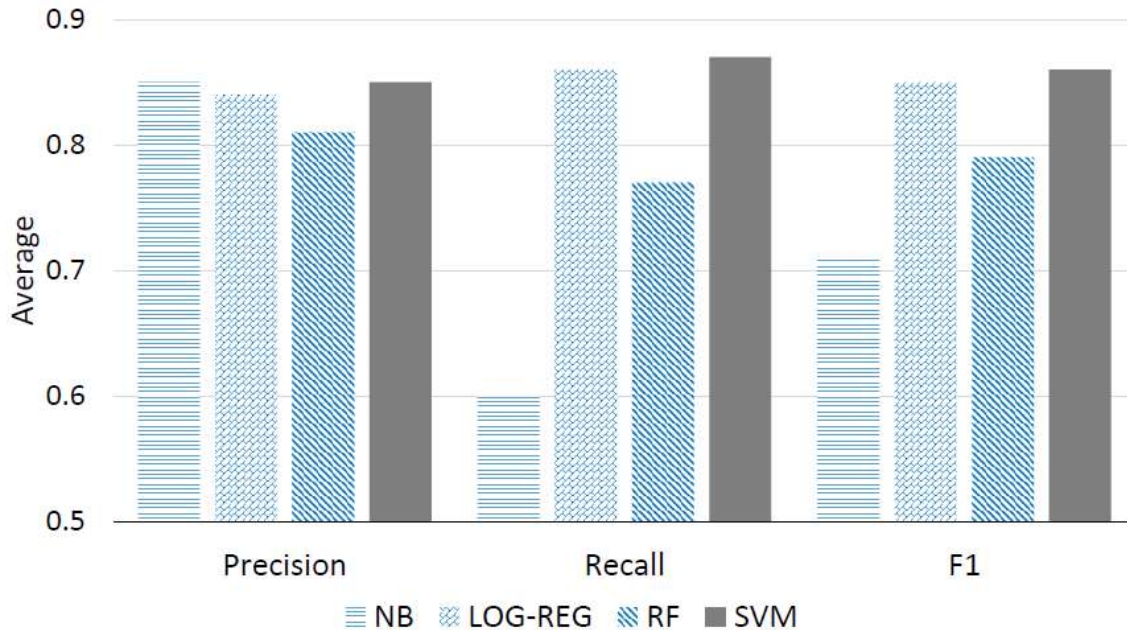
not expressly state it uses a Random Forest tree-based model.” Pet. at 67. Petitioner then cites then cites *Nunes* for the disclosure of a random forest tree-based model. *Id.* I disagree that *Nunes* discloses this claim element for multiple reasons.

219. First, for the reasons stated regarding claim 4, a person of ordinary skill in the art would not have been motivated to further combine the combination *Smyth* and *Phillips* with *Nunes*, because *Smyth* and *Phillips* do not analyze communication content, and because *Phillips* actually teaches away from that concept, in favor of analyzing solely communication metadata.

220. Second, *Nunes* merely references the phrase “Random Forst” as a “well-known classification technique,” but provides no further detail about its use. *Nunes* at 2. *Nunes* does not disclose the details of its implementation, and therefore does not teach or suggest an RF method “including multiple decision tree predictors applied in combination to classify the software vulnerability.” Petitioner does not claim that this claim element is inherent in RF methods, and indeed it is not. Other methodologies of implementing an RF method exist, for example logistic regression or Naive Bayes models.

221. Further, *Nunes* teaches away from RF. *See, e.g., Nunes* at 4 and Fig. 2 (“Results” section states “SVM with linear kernel performed the best, recalling 87% of the relevant products while maintaining a precision of 85% (Fig. 2). SVM performed the best likely due to the fact it maximizes generality as opposed to

minimizing error.”) This statement is supported by *Nunes*’ results shown in figure 2.



*Nunes* at Fig. 2. In all three instances, *Nunes* shows that SVM (dark gray) significantly outperforms RF(diagonal blue stripes).

#### ***D. Ground 4: Obviousness in light of Smyth, Phillips, and Kuperman***

222. Petitioner asserts that claim 6 is obvious in light of *Smyth*, *Phillips*, and *Kuperman*.

##### **1. Claim 6**

**Claim 6.** The method of claim 1, further comprising testing the classification models by applying additional training data and one or more algorithms and evaluation metrics to optimize the classification models until the classification models compute the likelihood of exploitation according to a predefined error rate.

223. Claim 6 depends on Claim 1 and is not disclosed by the combination of *Smyth*, *Phillips*, and *Kuperman* for the same reasons as stated in regard to ground 2, Claim 1.

***E. Ground 5: Obviousness in light of Smyth, Phillips, and Marin***

224. Petitioner asserts that claims 7 and 13 are obvious in light of *Smyth*, *Phillips*, and *Marin*.

**1. Claim 7**

**Claim 7.** The method of claim 1, further comprising vectorizing text features derived from the plurality of datasets using term frequency-inverse document frequency to create a vocabulary of associated words.

225. Claim 7 depends on Claim 1 and is not disclosed by the combination of *Smyth*, *Phillips*, and *Marin* for the same reasons as stated in regard to ground 2, Claim 1.

**2. Claim 13**

**Claim 13.** The method of claim 1, further comprising vectorizing text features derived from textual content of the plurality of datasets using a predetermined natural language process (NPL [sic]).

226. Claim 13 depends on Claim 1 and is not disclosed by the combination of *Smyth*, *Phillips*, and *Marin* for the same reasons as stated in regard to ground 2, Claim 1.

***F. Ground 6: Obviousness in light of Smyth, Phillips, and Bozorgi***

227. Petitioner asserts that claims 8 and 12 are obvious in light of *Smyth*, *Phillips*, and *Bozorgi*.

## 1. Claim 8

**Claim 8.** The method of claim 1, further comprising:

sorting vulnerabilities associated with the plurality of datasets according to time;

training the classification model using the training data, the training data defining a first subset of the plurality of datasets associated with a predetermined period of time; and

testing the classification model using a second subset of the plurality of datasets associated with the predetermined period of time.

228. Petitioner notes that “*Smyth* . . . does not sort and divide vulnerability data for training and testing based on time.” Pet. at 73. Instead, Petitioner relies on *Bozorgi* to disclose the claim element of “training data defining a first subset of the plurality of datasets associated with a predetermined period of time” and “testing the classification model using a second subset of the plurality of datasets associated with the predetermined period of time.” Specifically, Petitioner relies on *Bozorgi*’s Online Exploit Detection (Section 5.4). Pet. at 73.

229. The claim element explicitly requires a single predetermined time period from which two subsets of data are extracted. The training data is “a first subset of the plurality of datasets associated with a predetermined period of time,” and the testing data is “a second subset of the plurality of datasets associated with the predetermined period of time.” This configuration allows overlap between training data and testing data, since the claim does not impose any requirement beyond that

they are both subsets of datasets associated with the same predetermined period of time.

230. Petitioner cites to *Bozorgi*'s disclosure of "Online Exploit Detection." Pet. at 73. This disclosure, in contrast to the '897 patent, *Bozorgi* does not allow overlap, because it relies on training data from time interval  $t_0$ – $t_1$ , whereas testing data comes from timer interval  $t_1$ – $t_2$ . Thus, *Bozorgi* discloses training data that always precedes testing data, in order to avoid the temporal causality problem. *Bozorgi* at 5 ("In a real-world deployment, however, system administrators would train the classification known vulnerabilities [training set from time period  $t_0$ – $t_1$ ] to make predictions about new ones [testing set from time period  $t_1$ – $t_2$ ]."). That does not disclose the claim element where training and testing data come from the same dataset associated with the same time period. Thus, Petitioner's reference to *Bozorgi* does not disclose a set of training data that may overlap with a set of testing data as claimed.

## **2. Claim 12**

**Claim 12.** The method of claim 1, further comprising:

accessing one or more features indicative of temporal connections between at least two of:

- a time associated with discussion of a vulnerability by users posting hacking-related content at a web forum prior to disclosure to a public vulnerability database;

- a time associated with disclosure of the vulnerability to the public vulnerability database; and

a time associated with exploitation of the vulnerability as obtained through exploits data associated with real-world exploitation of the vulnerability.

231. Petitioner cites ground 2, claim element 1.c for the proposition that “*Smyth-Phillips* teaches deriving features from various data sources 100, including features from web forums containing hacking content.” Pet. at 74. I disagree for the reasons set forth in my response to that claim element. Specifically, neither *Smyth* nor *Phillips* disclose utilizing the content of communications related to hacking.

## **IX. CONCLUSION**

232. My opinions are subject to change or revision. I may acquire additional opinions regarding arguments Petitioner, or its expert, may present, information I may receive in the future, or additional work I may perform. Based on the analysis I have conducted and for the reasons set forth herein, I have preliminarily reached the conclusions and opinions in this declaration.

233. In the event of a hearing or trial in this matter, I may create and/or use various exhibits for the purposes of demonstrating my testimony as discussed in this declaration. I have not yet created, directed the creation of, or selected any demonstrative exhibits for this purpose.

234. For at least the reasons stated herein, it is my expert opinion that Petitioner’s references do not disclose or suggest all elements of any claim of the ’897 patent. Thus, I understand that Defendant’s expert witness fails to establish a

*prima facie* case of obviousness. My review of the references, the combinations of references, and the materials I have considered leave me with the opinion that the patent claims are valid over the prior art. In reaching my opinions, and in stating the basis of my opinions, I have provided representative reasons why the asserted claims are not disclosed by the cited references. I may not have provided every reason why the asserted claims are not invalid over the references (including combinations) as many are self-evidently trivial and make-weight. I therefore reserve the right to supplement my opinion and provide additional bases for my opinions, as necessary, particularly if Petitioner raises new arguments or interpretations, or if any other new information becomes available to me during this proceeding.