

**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

---

**PATENT OWNER'S PRELIMINARY RESPONSE**

## TABLE OF CONTENTS

|      |                                                        |    |
|------|--------------------------------------------------------|----|
| I.   | INTRODUCTION .....                                     | 1  |
| II.  | RELATED PROCEEDINGS .....                              | 1  |
| III. | SKYSONG BACKGROUND .....                               | 1  |
| IV.  | THE ASSERTED PRIOR ART .....                           | 2  |
| A.   | <i>Sabottke</i> (EX1004) .....                         | 2  |
| B.   | <i>Smyth</i> (EX1005) .....                            | 4  |
| C.   | <i>Phillips</i> (EX1006).....                          | 4  |
| D.   | <i>Nunes</i> (EX1007) .....                            | 5  |
| E.   | <i>Kuperman</i> (EX1008) .....                         | 6  |
| F.   | <i>Marin</i> (EX1009) .....                            | 7  |
| G.   | <i>Bozorgi</i> (EX1010) .....                          | 7  |
| V.   | THE '897 PATENT .....                                  | 8  |
| A.   | Technical Background.....                              | 8  |
| B.   | Prosecution History .....                              | 10 |
| C.   | Priority Date .....                                    | 11 |
| VI.  | LEVEL OF ORDINARY SKILL IN THE ART .....               | 12 |
| A.   | Petitioner's Identification Of The Relevant Skill..... | 12 |
| VII. | PATENT OWNER'S CLAIM CONSTRUCTIONS .....               | 13 |
| A.   | "Social Connections".....                              | 14 |
| B.   | "Disclosure Date" .....                                | 15 |
| C.   | "Exploitation Date" .....                              | 16 |

|       |                                                                                                                    |    |
|-------|--------------------------------------------------------------------------------------------------------------------|----|
| D.    | “Likelihood Of Exploitation” .....                                                                                 | 16 |
| VIII. | PETITIONER HAS FAILED TO DEMONSTRATE A REASONABLE<br>LIKELIHOOD OF PREVAILING AS TO ANY CHALLENGED CLAIM.<br>..... | 18 |
| A.    | GROUND 1: Obviousness in Light of <i>Sabottke</i> .....                                                            | 18 |
| 1.    | Claim 1 .....                                                                                                      | 19 |
| 2.    | Claim 3 .....                                                                                                      | 33 |
| 3.    | Claim 5 .....                                                                                                      | 34 |
| 4.    | Claim 9 .....                                                                                                      | 34 |
| 5.    | Claim 10 .....                                                                                                     | 34 |
| 6.    | Claim 14 .....                                                                                                     | 35 |
| 7.    | Claim 15 .....                                                                                                     | 36 |
| 8.    | Claim 18 .....                                                                                                     | 36 |
| 9.    | Claim 20 .....                                                                                                     | 37 |
| B.    | GROUND 2: Obviousness in Light of <i>Smyth</i> and <i>Phillips</i> .....                                           | 37 |
| 1.    | Claim 1 .....                                                                                                      | 38 |
| 2.    | Claim 2 .....                                                                                                      | 47 |
| 3.    | Claim 3 .....                                                                                                      | 48 |
| 4.    | Claim 5 .....                                                                                                      | 48 |
| 5.    | Claim 10 .....                                                                                                     | 48 |
| 6.    | Claim 11 .....                                                                                                     | 49 |
| 7.    | Claim 13 .....                                                                                                     | 49 |

|     |                                                                             |    |
|-----|-----------------------------------------------------------------------------|----|
| 8.  | Claim 14.....                                                               | 50 |
| 9.  | Claim 15.....                                                               | 51 |
| 10. | Claim 17.....                                                               | 51 |
| 11. | Claim 18.....                                                               | 52 |
| 12. | Claim 19.....                                                               | 53 |
| 13. | Claim 20.....                                                               | 53 |
| 14. | Motivation to Combine .....                                                 | 54 |
| C.  | GROUND 3: Obviousness in light of <i>Smyth, Phillips, and Nunes</i> .....   | 56 |
| 1.  | Claim 4.....                                                                | 56 |
| 2.  | Claim 16.....                                                               | 58 |
| D.  | GROUND 4: Obviousness in light of <i>Smyth, Phillips, and Kuperman</i> ..   | 60 |
| 1.  | Claim 6.....                                                                | 60 |
| E.  | GROUND 5: Obviousness in light of <i>Smyth, Phillips, and Marin</i> .....   | 60 |
| 1.  | Claim 7.....                                                                | 60 |
| 2.  | Claim 13.....                                                               | 61 |
| F.  | GROUND 6: Obviousness in light of <i>Smyth, Phillips, and Bozorgi</i> ..... | 61 |
| 1.  | Claim 8.....                                                                | 61 |
| 2.  | Claim 12.....                                                               | 63 |
| IX. | SECONDARY CONSIDERATIONS OF NON-OBVIOUSNESS .....                           | 64 |
| A.  | Commercial Success.....                                                     | 64 |
| B.  | Long-Felt Need and Government Support .....                                 | 66 |

|    |                             |    |
|----|-----------------------------|----|
| C. | Failure Of Others .....     | 66 |
| D. | Unexpected Results .....    | 67 |
| E. | Skepticism of Experts.....  | 67 |
| F. | Praise By Others .....      | 67 |
| G. | Copying By Competitors..... | 67 |
| X. | CONCLUSION.....             | 68 |

## TABLE OF AUTHORITIES

### CASES

|                                                                                           |    |
|-------------------------------------------------------------------------------------------|----|
| <i>Graham v. John Deere Co. of Kan. City</i> ,<br>383 U.S. 1 (1966).....                  | 64 |
| <i>Phillips v. AWH Corp.</i> ,<br>415 F.3d 1303 (Fed. Cir. 2005) ( <i>en banc</i> ) ..... | 13 |

## TABLE OF ABBREVIATIONS

|                        |                                                                                                                                                                                                                     |
|------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Patent Owner</b>    | Skysong Innovations, LLC (“Skysong”)                                                                                                                                                                                |
| <b>Petitioner</b>      | CrowdStrike, Inc. (“CrowdStrike”)                                                                                                                                                                                   |
| <b>PTAB</b>            | Patent Trial and Appeals Board                                                                                                                                                                                      |
| <b>’897 patent</b>     | U.S. Patent No. 11,892,897 (EX1001)                                                                                                                                                                                 |
| <b><i>Sabottke</i></b> | Carl Sabottke et al., <i>Vulnerability Disclosure in the Age of Social Media: Exploiting Twitter for Predicting Real-World Exploits</i> , Proceedings of the 24th USENIX Security Symposium (2015) (EX1004)         |
| <b><i>Smyth</i></b>    | U.S. Patent No. 10,264,009 to Smyth et al. (EX1005)                                                                                                                                                                 |
| <b><i>Phillips</i></b> | Elizabeth Phillips et al., <i>Extracting Social Structure from DarkWeb Forums</i> , SOTICS 2015: The Fifth International Conference on Social Media Technologies, Communication, and Informatics, (2015) (EX1006)   |
| <b><i>Nunes</i></b>    | Eric Nunes et al., <i>Darknet and Deepnet Mining for Proactive Cybersecurity Threat Intelligence</i> , Proceedings of the 2016 IEEE Conference on Intelligence and Security Informatics (EX1007)                    |
| <b><i>Kuperman</i></b> | U.S. Patent No. 10,652,254 to Kuperman et al. (EX1008)                                                                                                                                                              |
| <b><i>Marin</i></b>    | Ericsson Marin et al., <i>Product Offerings in Malicious Hacker Markets</i> , Proceedings of the 2016 IEEE Conference on Intelligence and Security Informatics (EX1009)                                             |
| <b><i>Bozorgi</i></b>  | Mehran Bozorgi et al., <i>Beyond Heuristics: Learning to Classify Vulnerabilities and Predict Exploits</i> , Proceedings of the Sixteenth ACM Conference on Knowledge Discovery and Data Mining (KDD 2010) (EX1010) |

## LIST OF EXHIBITS

| Exhibit No. | Descriptions                                                                                                                                                                                                                                   |
|-------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2001        | Skysong's First Amended Complaint against CrowdStrike                                                                                                                                                                                          |
| 2002        | Declaration of Dr. Paulo Shakarian regarding discretionary denial with Appendix                                                                                                                                                                |
| 2003        | Civil Docket For Case # 7:25-cv-00040-DC-DTG                                                                                                                                                                                                   |
| 2004        | <i>Skysong v. CrowdStrike</i> , Dkt. No. 55, Scheduling Order                                                                                                                                                                                  |
| 2005        | Table of CrowdStrike IPRs and Parallel District Litigation Schedules                                                                                                                                                                           |
| 2006        | Civil Docket For Case # 2:25-cv-00098-JRG                                                                                                                                                                                                      |
| 2007        | Executive Order No. 14,028                                                                                                                                                                                                                     |
| 2008        | Declaration of Dr. Nenad Medvidovic                                                                                                                                                                                                            |
| 2009        | Joseph R. Zipkin et al., <i>Predicting Exploitation of Disclosed Software Vulnerabilities Using Open-Source Data</i> , Proceedings of the 3rd ACM on International Workshop on Security And Privacy Analytics (2017) 45–53 (“ <i>Zipkin</i> ”) |
| 2010        | Haipeng Chen, et al., <i>Using Twitter to Predict When Vulnerabilities will be Exploited</i> , KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Pages 3143 – 3152 (“ <i>Chen</i> ”)  |
| 2011        | Second Declaration of Dr. Paulo Shakarian regarding the preliminary response                                                                                                                                                                   |
| 2012        | CYR3CON Product Roadmap                                                                                                                                                                                                                        |
| 2013        | Second Declaration of Dr. Paulo Shakarian regarding the preliminary response redacted version                                                                                                                                                  |
| 2014        | Patent Owner's Preliminary Response redacted version                                                                                                                                                                                           |

## LIST OF CHALLENGED CLAIMS

| <b>Limitation</b> | <b>Claim Language</b>                                                                                                                                                                                              |
|-------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Claim 1</b>    |                                                                                                                                                                                                                    |
| 1                 | A method for assessing a likelihood of exploitation of software vulnerabilities, comprising:                                                                                                                       |
| 1.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:                                             |
| 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;                                         |
| 1.c               | accessing features from the plurality of datasets that include measures computed from social connections of users posting hacking-related content                                                                  |
| 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;      |
| 1.e               | accessing one or more features associated with a software vulnerability; and                                                                                                                                       |
| 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,           |
| 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.                     |
| <b>Claim 2</b>    |                                                                                                                                                                                                                    |
| 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.                                                    |
| <b>Claim 3</b>    |                                                                                                                                                                                                                    |
| 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. |

| <b>Claim 4</b>  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 4               | The method of claim 3, further comprising: <ul style="list-style-type: none"> <li>- aligning the exploits data with the vulnerability data;</li> <li>- and cleaning the exploits data of noise and predetermined portions of the exploits data that is irrelevant to associated software vulnerabilities.</li> </ul>                                                                                                                                                                                                                  |
| <b>Claim 5</b>  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 5               | The method of claim 1, wherein certain features correspond to a known vulnerability obtained from the plurality of datasets.                                                                                                                                                                                                                                                                                                                                                                                                          |
| <b>Claim 6</b>  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 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.                                                                                                                                                                                                                 |
| <b>Claim 7</b>  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 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.                                                                                                                                                                                                                                                                                                                                |
| <b>Claim 8</b>  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 8               | The method of claim 1, further comprising: <ul style="list-style-type: none"> <li>- sorting vulnerabilities associated with the plurality of datasets according to time;</li> <li>- 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;</li> <li>- and testing the classification model using a second subset of the plurality of datasets associated with the predetermined period of time.</li> </ul> |
| <b>Claim 9</b>  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 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.                                                                                                                                                                                                                                                                                                                                              |
| <b>Claim 10</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 10              | The method of claim 1, further comprising detecting, from the plurality of datasets, vulnerabilities that appear before an associated exploitation date.                                                                                                                                                                                                                                                                                                                                                                              |

| <b>Claim 11</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 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.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Claim 12</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 12              | <p>The method of claim 1, further comprising accessing one or more features indicative of temporal connections between at least two of:</p> <ul style="list-style-type: none"> <li>- 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;</li> <li>- a time associated with disclosure of the vulnerability to the public vulnerability database; and</li> <li>- a time associated with exploitation of the vulnerability as obtained through exploits data associated with real-world exploitation of the vulnerability.</li> </ul>                                                                      |
| <b>Claim 13</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 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]).                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| <b>Claim 14</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 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. |
| <b>Claim 15</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 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.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

| <b>Claim 16</b> |                                                                                                                                                                                                                                                                                                                                                                              |
|-----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 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.                                                                                                      |
| <b>Claim 17</b> |                                                                                                                                                                                                                                                                                                                                                                              |
| 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.                                                                |
| <b>Claim 18</b> |                                                                                                                                                                                                                                                                                                                                                                              |
| 18              | A method of prioritizing vulnerabilities using cyber threat intelligence, comprising:                                                                                                                                                                                                                                                                                        |
| 18.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:                                                                                                                                                                                                        |
| 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;                                                                                                                                                                               |
| 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;                                                                                                                                                                                                  |
| 18.d            | accessing one or more features associated with a software vulnerability; and                                                                                                                                                                                                                                                                                                 |
| 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. |
| <b>Claim 19</b> |                                                                                                                                                                                                                                                                                                                                                                              |
| 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.                                                                                                                                                                                                             |

| <b>Claim 20</b> |                                                                                                                                                                                                                     |
|-----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 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. |

## **I. INTRODUCTION**

Patent Owner Skysong respectfully submits this Preliminary Response to the Petition for *Inter Partes* Review dated July 25, 2025 (“Petition”) of U.S. Patent No. 11,892,89 (EX1001, “’897 patent”). This response is timely filed. *See* Paper 4.

## **II. RELATED PROCEEDINGS**

The ’897 patent is the subject of the following related litigation:

- *Skysong Innovations, LLC v. CrowdStrike, Inc. and CrowdStrike Holdings, Inc.*, Case No. 7:25-cv-00040, filed January 31, 2025 in the United States District Court for the Western District of Texas; and
- *Skysong Innovations, LLC v. Fortinet, Inc.*, Case No. 2:25-cv-00098, filed January 31, 2025, in the United States District Court for the Eastern District of Texas.

## **III. SKYSONG BACKGROUND**

Skysong is a nonprofit organization, and the exclusive intellectual property management and technology transfer organization for Arizona State University (“ASU”). Skysong works closely with ASU and its technology industry partners to translate ASU’s innovations into commercial applications, fostering economic development and public benefit. *See* EX2001 (Skysong’s First Amended Complaint), ¶¶3–4. Skysong significantly contributes to the mission statement of

the University and has become a core component of Arizona’s innovation ecosystem and a driver of high-tech industry in the state.

The ’897 patent stems from research conducted by Dr. Paulo Shakarian and his team. *See* EX2002, ¶¶5–6, 11. It discloses artificial intelligence-based methods for detecting and mitigating cybersecurity threats, including analyzing data from online sources to predict and respond to vulnerabilities. *See* EX1001, 3:26–4:27.

#### **IV. THE ASSERTED PRIOR ART**

Dr. Nenad Medvidovic, Department Chair of the Computer Science Department of the University of Southern California and leading artificial intelligence and cybersecurity authority, has been retained by Skysong as an independent expert to assist in the technical analysis of the ’897 patent and the asserted prior art. He has provided a declaration (EX2008, “*Declaration Of Nenad Medvidovic, Ph. D., In Support Of Patent Owner’s Preliminary Response*”) setting forth his opinions, including a synopsis of the relevant aspects of the asserted prior art. His synopsis is relevant and reproduced here for reference:

##### **A. Sabottke (EX1004)<sup>1</sup>**

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

---

<sup>1</sup> EX2008 at ¶¶79–84 (*Sabottke*).

to be exploited in real-world attacks. EX1004 at 2. Notably, *Sabottke* relies exclusively on Twitter for its input data.

*Sabottke* does not consider “the detection of zero day attacks” and admits that “a complete system for early exploit detection . . . likely requires mechanisms beyond the realm of Twitter analytics.” *Sabottke* at 3. *Sabottke* further explicitly identifies “data scarcity” as a significant challenge to its model:

**Data scarcity.** [O]nly a small subset of Twitter users discuss vulnerability exploits (approximately 32,000 users), and they do not always mention the CVE 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 than 50 associated tweets. This data scarcity compounds the challenge of class imbalance for reducing false positives and false negatives.

*Id.* at 4. CVE numbers are “Common Vulnerabilities and Exposures” numerical identifiers first published in the National Vulnerability Database (NVD). ’897 patent at 2:28–29.

In summary, *Sabottke*’s training data is restricted to NVD, Open Source Vulnerability Database (OSVDB), and Twitter—wherein Twitter is the only

dataset including hacker data. *Sabottke* only considers vulnerabilities after the disclosure date. The social information on Twitter is used directly and is not further analyzed with other methods. Further, *Sabottke*'s training data includes information that occurred after exploitation in the wild. Additionally, 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 binary classes: “exploited” or “non-exploited.” No further indication of any “likelihood of exploitation” is produced.

**B. *Smyth* (EX1005)<sup>2</sup>**

*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*.” *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.

**C. *Phillips* (EX1006)<sup>3</sup>**

*Phillips* refers to Elizabeth Phillips et al., *Extracting Social Structure from DarkWeb Forums*, SOTICS 2015: The Fifth International Conference on Social

---

<sup>2</sup> EX2008 at ¶¶85–88 (*Smyth*).

<sup>3</sup> EX2008 at ¶¶89–91 (*Phillips*).

Media Technologies, Communication, and Informatics, (2015) (EX1006, “*Phillips*”).

*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.*

*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)<sup>4</sup>**

*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.

---

<sup>4</sup> EX2008 at ¶¶92–94 (*Nunes*).

*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. Its sole focus is on identifying products (from market data) and communications (from forum data) relevant to potential threats. *Nunes* at 2.

**E. *Kuperman* (EX1008)<sup>5</sup>**

*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*.” *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.*

*Kuperman* discloses that the WAF may utilize an Artificial Neural Network 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).

---

<sup>5</sup> EX2008 at ¶¶95–97 (*Kuperman*).

**F. *Marin* (EX1009)<sup>6</sup>**

*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*”).

*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)<sup>7</sup>**

*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*”).

*Bozorgi* discloses an exploit prediction system relying 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

---

<sup>6</sup> EX2008 at ¶¶98–99 (*Marin*).

<sup>7</sup> EX2008 at ¶¶100–101 (*Bozorgi*).

extract from the text fields, time stamps, cross-references, and other entries in existing vulnerability disclosure reports.” *Id.*

## V. THE '897 PATENT

### A. Technical Background

Dr. Medvidovic provides a helpful explanation of the technical background of the '897 patent in paragraphs 66–75 of his declaration, reproduced in full here:

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”.

EX2008 at ¶¶66–75.

## **B. Prosecution History**

The provisional application that resulted in the '897 patent was filed on Nov. 3, 2017, and the corresponding PCT application on Oct. 26, 2018. EX1001. An office action issued on September 30, 2022, included a preliminary rejection of claims 1–9 and 15 (claims 10–14 had been withdrawn). EX1002 at 349. Patent Owner responded on December 30, 2022, and amended its claims and added claims 16–23. EX1002 at 396. A final rejection issued on April 13, 2023, rejecting claims 1–9, 15–16 and 20–23, and objecting to claims 17–19. EX1002 at 454. Patent Owner responded on July 13, 2023, and explicitly added the limitation requiring “features from the plurality of datasets that include measures computed from social connections of users posting hacking-related content” and “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.” EX1002 at 471, 472 (emphasis added). In light of the amendment, the examiner issued a notice of allowance on September 28, 2023, noting that “Applicant’s amendments . . . overcome the prior art.” EX1002 at 522.

### **C. Priority Date**

The ’897 patent derives priority from provisional application No. 62/581,123, filed on Nov. 3, 2017. The provisional application fully discloses the subject matter of the claims in the ’897 patent. EX2008 at ¶77. Petitioner alleges the following priority dates for the asserted prior art:

- 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)

Since these dates predate the provisional patent application as well, Patent Owner’s analysis relies on the specification of the ’897 patent itself. However, Patent Owner is still investigating the veracity of Petitioner’s alleged dates. To the

extent these dates are inaccurate and in fact post-date the provisional application, Patent Owner reserves the right to identify the respective portions of the provisional application supporting the claims of the '897 patent.

## **VI. LEVEL OF ORDINARY SKILL IN THE ART**

In consideration of 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 in reliance on Dr. Medvidovic's expert opinion, Patent Owner defines the relevant skill in the art as follows: A person of ordinary skill in the art ("POSITA") at the time of the filing of the '897 patent would have at least a Bachelor's degree in computer science ("CS"), data science / engineering, or cybersecurity, and 2–4 years of experience working in the field of software development, artificial intelligence ("AI"), machine learning ("ML"), data mining, or cybersecurity. EX2008 at ¶34. A graduate degree in CS, AI, data science/engineering, cybersecurity, or software engineering with specific courses in cybersecurity may be substituted wholly or in part for the relevant work experience. *Id.* Dr. Medvidovic qualifies as a POSITA. *Id.* at ¶35.

### **A. Petitioner's Identification Of The Relevant Skill**

Petitioner notes a different level skill: "A POSITA at the time of the '897 patent would have had a bachelor's degree in electrical or computer engineering,

computer science, or the equivalent, and (a) either a master’s degree in electrical or computer engineering, computer science, or the equivalent, or (b) two years of work/research experience in computer systems and networks.” Pet. at 3–4.

Petitioner relies on the expert declaration of Dr. Nicholas Bambos (EX1003). Dr. Bambos identifies his education as a Ph.D. in Electrical Engineering and Computer Sciences, a M.S. in EECS and a M.A. in Mathematics, and a degree in Electrical Engineering from the National Technical University of Athens-Greece. EX1003 at ¶5. He is a professor in the school of Engineering, the school of Electrical Engineering, and the school of Management Science & Engineering. *Id.* at ¶6.

Patent Owner disagrees with this level of skill, for lack of any requirement of experience in the field of artificial intelligence, machine learning, or cybersecurity. Dr. Bambos’ experience appears to be in the “Computer Systems & Networks Performance Engineering” area with focus on “high performance designs of computer systems and networks.” His declaration and curriculum vitae attached thereto do not provide any detail regarding his experience, if any, in the relevant field of artificial intelligence, machine learning, or cybersecurity.

## **VII. PATENT OWNER’S CLAIM CONSTRUCTIONS**

Patent claims are construed using the framework set forth in *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–15 (Fed. Cir. 2005) (*en banc*). CrowdStrike relies on

the plain and ordinary meaning of all claim terms. Pet. at 5 and Bambos at ¶36.

Patent Owner agrees but prefers to identify what that plain and ordinary meaning is to avoid any ambiguities.

**A. “Social Connections”**

A POSITA would understand the term “social connections” in the ’897 patent as including “the in-degree and out-degree a given user has to other active users in the social network.” EX2008 at ¶103.

The ’897 patent sets forth the details as to how this data is derived. First, the ’897 patent extracts and analyzes data from various darkweb and deepweb (“DW”) sources. The invention uses this data to identify “a set of features computed from the social network structure of users posting in a forum.” EX2008 at ¶104 (citing ’897 patent at 11:18–22).

This set of features contains 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.

EX2008 at ¶105 (citing '897 patent at 11:29–38 and 21:6–47).

These social features and particularly “social connection” metrics can be used to determine “Importance Measures.” EX2008 at ¶106 (citing '897 patent at 21:31–46. These “incorporate the ‘centrality of users’ in the relevant social metric to assign weight to communication to or from that user.” EX2008 at ¶106. The “centrality of users” is determined using the users’ in-degree and out-degree, which the '897 patent defines as the number of edges pointing to the user (in-degree) and the number of edges originated from the user (out-degree). EX2008 at ¶107 (citing '897 patent at 21:12–14).

Thus, a POSITA would understand “social connections” as including the “in-degree and out-degree a given user has to other active users in the social network.”

## **B. “Disclosure Date”**

A POSITA would understand the term “disclosure date” as covering the day the National Institute of Standards and Technology (“NIST”) officially releases the vulnerability number to the public. EX2008 at ¶109 (citing '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”**

A POSITA would understand the term “exploitation date” of a vulnerability as covering the date the exploit was first detected in the wild. EX2008 at ¶110 (citing ’897 patent at 9:42–43 (“the exploitation date of a vulnerability was defined as the date it was first detected in the wild.”)).

### **D. “Likelihood Of Exploitation”**

A POSITA would understand the term “likelihood of exploitation” as requiring a gradient in order to facilitate “ranking vulnerabilities” that are likely to be exploited, which would be impossible in a binary classification model consisting only of “exploited” and “not exploited.” EX2008 at ¶¶111, 113 (citing ’897 patent at 1:36–38, 7:15–17). The ’897 patent discloses in multiple instances the ability to “increase” a likelihood of exploitation, and even assigns precise percentage values in certain examples. ’897 patent at 4:5–7.

For further examples, the ’897 patent teaches that the “presence of proof-of-concept exploits in EDB increases the likelihood of a vulnerability being exploited.” ’897 patent at 11:1–3; *see also 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.”); ’897 patent at 1:36–38 (once vulnerabilities are disclosed publicly, “the likelihood of those vulnerabilities being exploited increases.”); ’897 patent at

4:1–4 (vulnerability mentions on EDB (Exploit Database), ZDI (Zero Day Initiative), and DW (Dark Web) further “increase in the vulnerability exploitation likelihood over vulnerabilities only disclosed on NVD.”). A POSITA would therefore understand that an “increase” in likelihood of exploitation means something akin in the increase from “low likelihood” to “medium likelihood” or to “high likelihood.” EX2008 at ¶111.

This is further confirmed by Table 6, which indicates different likelihoods for vulnerabilities disclosed on different sources.

TABLE 6

| Precision, Recall, F1 measure for vulnerabilities mentioned on DW, ZDI, and EDB. |           |           |        |            |
|----------------------------------------------------------------------------------|-----------|-----------|--------|------------|
| Source                                                                           | Case      | Precision | Recall | F1 measure |
| DW                                                                               | NVD       | 0.23      | 0.38   | 0.27       |
|                                                                                  | NVD + DW  | 0.67      | 0.375  | 0.48       |
| ZDI                                                                              | NVD       | 0.16      | 0.54   | 0.25       |
|                                                                                  | NVD + ZDI | 0.49      | 0.24   | 0.32       |
| EDB                                                                              | NVD       | 0.15      | 0.56   | 0.24       |
|                                                                                  | NVD + EDB | 0.31      | 0.40   | 0.35       |

’897 patent, Table 6. Using its precise data, 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.

Thus, the '897 patent uses a likelihood spectrum, not a mere binary choice, and the term “likelihood of exploitation” requires multiple different likelihoods of exploitation beyond merely the binary choice of “exploited” or “not exploited.”

### **VIII. PETITIONER HAS FAILED TO DEMONSTRATE A REASONABLE LIKELIHOOD OF PREVAILING AS TO ANY CHALLENGED CLAIM.**

As a threshold matter, five of Petitioner’s seven references were available to the examiner during prosecution, and the first primary references, *Sabottke*, was explicitly addressed and distinguished in the specification. Each of the six grounds raised by petitioner relies on one or more of these previously considered references. In other words, Petitioner brings no ground that presents only new art.

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

Petitioner’s first ground asserts that claims 1, 3, 5, 9–10, 14–15, 18, and 20 are obvious in light of *Sabottke*.

As a threshold matter, 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:

For the following reasons, claim 1 is not taught or suggested by *Sabottke*.

*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;*

### (1) No Plurality Of Datasets Comprising Hacker Communications

*Sabottke* exclusively relies on Twitter as its sole source of hacker communications. In addition to Twitter, Petitioner also identifies “the NVD and OSVDB databases” as the “plurality of datasets from predetermined data sources.” Petition at 18, 19. This cannot be correct because the datasets must comprise “hacker communications.” This is explicitly required by the claim language.

Claim 1 requires that “the plurality of datasets include[es] training data comprising hacker communications.” ’897 patent, claim 1 (emphasis added). In contrast, claim 18 requires only that “at least a portion of the plurality of datasets defin[es] training data including hacker communications.” *Id.*, claim 18 (emphasis added). Thus, while claim 18 only requires that a portion of the plurality of datasets comprise hacker communications, claim 1 imposes that requirement on the entire plurality of datasets.

Dr. Medvidovic confirms that in *Sabottke* “Twitter is the only [dataset] that may comprise hacker communications.” EX2008 at ¶120. “Neither NVD (National Vulnerability Database) nor OSVDB (Open Source Vulnerability Database) comprise hacker communications. Both are vulnerability databases that publish comprehensive lists of vulnerabilities (NVD) and additional attributes.” *Id.*

*Sabottke* confirms that the only data it obtains from these supplemental databases is vulnerability context. *Id.* at 121 (*Sabottke* “only derives vulnerability information from NVD and OSVDB.”); *see also Sabottke* at 3, fn. 7; at 5 (“To obtain context about the vulnerabilities discussed on Twitter, we query the National Vulnerability Database (NVD) . . . we crawl the Open Sourced Vulnerability Database (OSVDB) [9] for a few additional attributes, including the disclosure dates and categories of the vulnerabilities . . .”) (emphasis added). Thus, *Sabottke* confirms that these supplemental sources (NVD and OSVDB) provide only contextual information about vulnerabilities, and do not contain hacker communications. The ’897 patent confirms that NVD comprises only vulnerability information. ’897 patent at 1:60–64 (“NIST provides the National Vulnerability Database (NVD) which comprises of a comprehensive list of vulnerabilities disclosed”).

This is explicitly confirmed by *Sabottke*, noting that the advantage of using Twitter 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 (emphasis added). Instead, *Sabottke* only uses NVD and OSVDB to gather information about the vulnerabilities discussed in the Twitter dataset. EX2008 at ¶124. According to *Sabottke*:

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. *Sabottke* thus gathers “hacker communications” exclusively from a single dataset gathered from Twitter and confirms that the NVD and OSVDB datasets do not comprise hacker communications. EX2008 at ¶125 (*Sabottke* queries NVD only for “CVSS scores, products affected, and additional references about these vulnerabilities,” and OSVDB only for “disclosure dates and categories of vulnerabilities.”).

“In contrast, the ’897 patent gathers ‘hacker communications’ from a much larger and diverse dataset.” EX2008 at ¶126. As Dr. Medvidovic explains, the ’897 patent discloses “a database of 46 darkweb and deepweb forums.” *Id.* (citing ’897 patent at 9:50–53). “This comes with distinct advantages.” *Id.* (citing ’897 patent at 12:5–6 (“Higher variability between [hacker communication] data sources 110 implies higher information gain and less redundancy.”)).

Therefore, *Sabottke* does not disclose “a plurality of datasets . . . including training data comprising hacker communications.” Instead, *Sabottke* discloses only a single dataset—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 currently is the best one.’” EX2008 at ¶127 (citing ’897 patent at 15:42–46).

Dr. Medvidovic cites an additional corroborating source, which concludes 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.” EX2008 at ¶127 (citing 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 (EX2009, “*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***

*(1) No Social Connections*

As set forth 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”).

*Sabottke* does not disclose—and Petitioner does not identify any element of *Sabottke* that discloses—“social connections.” An analysis of “social connection” “would require *Sabottke* to analyze the structure of the social network of the authors of the Twitter data it utilizes.” EX2008 at ¶129. Dr. Medvidovic explains that “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.” *Id.* *Sabottke* has no such disclosure and does not consider the “social connection” of its Twitter authors. *Id.*

Dr. Medvidovic also confirms that the analysis of social connections, as claimed in the ’897 patent, would require “a sophisticated analysis of Twitter author’s social networks.” *Id.* at ¶130. This would require “significantly deeper data collection and analysis than what is proposed by *Sabottke*.” *Id.* Thus, 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.

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 EX1003 at ¶63). However, "none of these disclosures identify or disclose 'social connections.'" EX2008 at ¶131. Dr. Medvidovic explains:

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.

EX2008 at ¶132.

Petitioner’s expert, Dr. Bambos, further cites *Benjamin* (EX1013) to disclose “social connections.” EX1003 at ¶63 (citing *Benjamin* at 7–8). However, *Benjamin* (entitled *Securing Cyberspace: Identifying Key Actors In Hacker Communities*), focuses only on identifying the most important actors in a community, and does not reference “social connections” within hacker social networks.

Instead, *Benjamin* explored two major hacker web forums: one from the United States (Hackhound.org) and one from China (Unpack.cn). *Benjamin* at 5–6. Both communities feature a peer-assigned reputation score system and allow users to attach hacking tools or program source code to their messages. *Benjamin* at 6. The study tested six features across two categories: (1) Average Message Length, (2) Number Of Replies Per Thread, (3) Number Of Threads Involved, (4) Tenure, (5) Sum Of Attachments (files/code), and (6) Total Messages. *Benjamin* at 7. “Social connections” to other users (i.e., in-edges and out-edges) were not considered. Thus, *Benjamin* also does not disclose use of “social connections.”

Notably, the “social connections” claim element was added during prosecution and was explicitly relied upon by the examiner to allow this claim. EX1002 at 471, 472 (amendment) and 517-523 (notice of allowance).

Therefore, *Sabottke* does not disclose or suggest 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;***

***(1) No Likelihood of Exploitation As Claimed***

As set forth above, “likelihood of exploitation” in the ’897 patent refers to a likelihood gradient in order to facilitate “ranking vulnerabilities” that are likely to be exploited. *See* sec. VII.D. above (construction of “likelihood of exploitation”). This would be impossible in a binary classification model consisting only of “exploited” and “not exploited.” EX2008 at ¶113 (citing ’897 patent at 7:15–17).

Petitioner argues 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. “This binary choice presents only information regarding whether a given vulnerability is expected to be exploited or not.” EX2008 at ¶135. “It does not include “a likelihood of exploitation,” which instead would require a classification such as “low”, “medium”, or “high” to indicate a probability.” *Id.* This is confirmed by the patent, which aims to “prioritize patching.” ’897 patent at abstract and 1:24–28. As Dr. Medvidovic notes, “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.” EX2008 at ¶135.

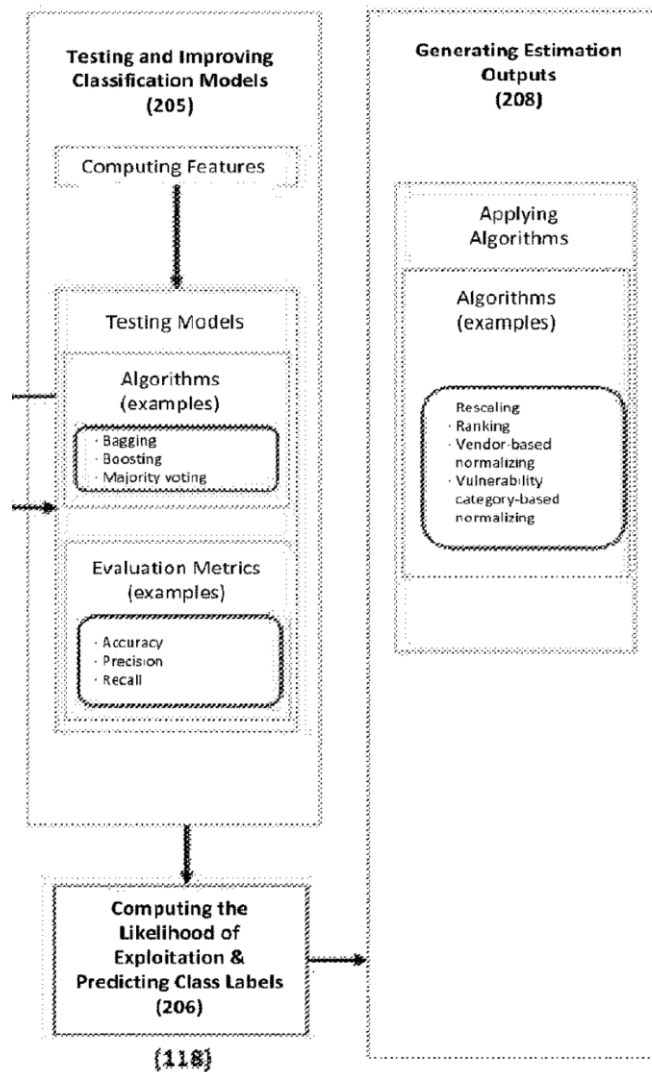
The specification of the ’897 patent confirms that “likelihood of exploitation” is a gradient. *See, e.g.*, ’897 patent, table 6 (mention on NVD increases the likelihood of exploitation); *id.* at 4:1–4 (vulnerability mentions on EDB, ZDI, and DW further “increase in the vulnerability exploitation likelihood over vulnerabilities only disclosed on NVD.”). “Thus, the ’897 patent uses a likelihood spectrum, not a binary choice.” EX2008 at ¶140. The ’897 patent even quantifies the stepwise increase. ’897 patent at 4:5–7 (“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%”).

Figure 13 of the ’897 patent further corroborates this. ’897 patent, Fig. 13. The “Likelihood of Exploitation” is computed in block 206, which is fed to Generating Estimation Outputs 208. *Id.* Dr. Medvidovic explains this in detail:

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.

EX2008 at ¶136 (citing '897 patent at 11:1–14). Therefore, “*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).

For these reasons, *Sabottke*'s binary classification of "exploited" and "not exploited" does not disclose or suggest 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,***

***(1) No Likelihood of Exploitation As Claimed***

As stated regarding claim element 1.d, *Sabottke* does not disclose computing the claimed class label defining a likelihood of exploitation. Instead, "all vulnerabilities with the class label "exploited" are treated equally by *Sabottke*, and no likelihood value is assigned." EX2008 at ¶138.

***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.***

***(1) No Predicting "An Actual Exploitation . . . Before Disclosure"***

As noted by Dr. Medvidovic, "*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.'" EX2008 at 142 (citing *Sabottke* at 3). "Thus, *Sabottke* does not disclose predicting 'an actual exploitation of the respective software vulnerabilities before disclosure" (emphasis added).'" *Id.*

Contrary to this explicit limitation, 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). But this statement only 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.” It does not relate to predicting actual exploitation before disclosure. “*Sabottke*’s reference to identifying features ‘before disclosure’ does not support that it predicts actual exploitation ‘before disclosure.’” EX2008 at ¶143.

Petitioner further argues that *Sabottke*’s identified “47 vulnerabilities that were mentioned in the Twitter stream before their public disclosure.” Pet. at 26 (citing *Sabottke* at 9). This statement also does not support pre-disclosure exploitation. Instead, it only tracks Twitter traffic regarding a vulnerability before disclosure thereof. EX2008 at ¶144. *Sabottke* then “investigate these cases manually to determine the sources of this information.” *Id.* (citing *Sabottke* at 9). This does not support exploit prediction before disclosure.

Finally, Petitioner argues that *Sabottke* acknowledges that “[w]hile the first tweets about a vulnerability precede the creation of IPS [Intrusion Detection System] or AV [Antivirus] 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). But that statement only confirms that *Sabottke* does not teach pre-disclosure exploit prediction. EX2008 at ¶145.

*Sabottke*’s further discussion confirms that this claim element is not disclosed. Specifically, since “these first tweets are typically not informative enough”—*Sabottke* 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. As Dr. Medvidovic explains:

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.

EX2008 at ¶146. This is corroborated by 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 (EX2010, “*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.”).

Notably, *Sabottke*’s simulation can arbitrarily modify the time of detection after the first tweet by lowering its precision threshold. *Sabottke* at 13 (“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.” EX2008 at ¶147. “Instead, the random data set *Sabottke* uses to train its classifiers include tweets both before and after disclosure.” *Id.* Therefore, *Sabottke* does not disclose 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*.” EX2008 at ¶147 (emphasis in original).

The ’897 patent specifically identifies this as an undesirably error: 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 (“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.”).

“Since *Sabottke*’s classifiers are not available *before* disclosure (they are created only after disclosure, based in part on post-disclosure tweets), *Sabottke* does not disclose pre-disclosure prediction.” EX2008 at ¶149.

Notably, the “before disclosure” claim element was added during prosecution and was explicitly relied upon by the examiner to allow this claim. EX1002 at 471, 472 (amendment) and 517-523 (notice of allowance).

## **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.

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.

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.

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.

As set forth above, the '897 patent defines the “exploitation date” of a vulnerability as the date the exploit was first detected in the wild. *See* sec. VII.C. above (construction of “exploitation date”); *see also* '897 patent at 9:42–43.

“*Sabottke* is incapable of meeting and does not disclose this requirement.”

EX2008 at ¶153.

*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. Therefore, *Sabottke* does not disclose this claim element which requires 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.

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. Patent Owner therefore incorporates its corresponding response to claim 1. *See* section VIII.A.1.

Petitioner specifically 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. Patent Owner therefore incorporates its corresponding response to these claim elements.

For the reasons explained regarding claim 1, *Sabottke* does not disclose at least the following elements: “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.g).

## **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.

Petitioner cites 1.f regarding this claim. Pet. at 32. Patent Owner therefore incorporates its analysis of that claim element. 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;*

Petitioner addresses this claim element only by reference to claim element 1.d. Patent Owner therefore incorporates its 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.***

Petitioner addresses this claim element only by reference to claim element 1.f and 1.f.1. Patent Owner therefore incorporates its analysis of that claim element. 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.

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***

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

As a threshold matter, the '897 patent cites *Phillips*. EX1002 at 173 (IDS listing *Phillips*); *see also* EX1001 at 2.

(1) *Phillips Is Not Analogous Art*

*Phillips* is not analogous art. *Phillips* focuses on extracting social structures from darkweb forums and adds no predictive modeling. 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. To accomplish this goal, *Phillips* relies exclusively on communication metadata. *Id.* Indeed, *Phillips* intentionally did not rely on “state-of-the-art Computational Linguistic techniques” to analyze communication content. *Id.* (“Whilst existing work attempts to identify these structures through state of-the-art Computational Linguistic techniques, our work relies on the communication metadata alone.”). *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 when evaluating millions of messages.”).

**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;***

*(1) No Hacker Communications*

Petitioner argues that *Smyth*'s "data sources 110," which comprise "online discussions 240" including "on social media and discussion forums," disclose this claim element. Pet. at 38. However, Petitioner recognizes that "*Smyth* does not expressly discuss accessing 'hacker communications' in those forums." Pet. at 38. Petitioner then turns to *Phillips* for the claimed "hacker communications." Pet. at 38. But *Phillips* explicitly teaches away from analyzing communication content, and instead relies on metadata alone. *Phillips* at 97 ("discounting "state-of-the-art Computational Linguistic techniques" in favor of "communication metadata alone.").

First, *Phillips* is not analogous art. *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. *Phillips* does not address cyber security vulnerabilities or exploits at all. Further, it disregards the content of the communications and relies exclusively on communication metadata. *Id.*

Second, *Phillips* teaches away from utilizing the content of communications. *Phillips* at 97. Despite being familiar with "state-of-the-art Computational

Linguistic techniques” for analyzing communication content, *Phillips* intentionally did not do so. *Id.* (“Whilst existing work attempts to identify these structures through state-of-the-art Computational Linguistic techniques, our work relies on the communication metadata alone.”). Dr. Medvidovic explains that “that is because use of communication content—*i.e.*, ‘hacker communications’—is a complex task.” EX2008 at ¶168. *Phillips* specifically notes the difficulty of scaling a system that relies on 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.”). “Thus, *Phillips* does not collect or consider ‘hacker communications.’” EX2008 at ¶168. “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.” *Id.*

For these reasons, the combination of *Smyth* and *Phillips* does not teach or suggest this claim element. Further, because *Phillips* is not analogous art, and because *Phillips* teaches away from utilizing the content of communications (as detailed further in the “Motivation to Combine” section below (see Section VIII.B.14), 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***

***(1) No Social Connections***

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. *Phillips* does not disclose this concept either.

*Phillips* discloses using “Social Network Analysis (SNA) techniques” to analyze dark web forums. However, *Phillips* does not compute or refer to “social connections” of hackers. Petitioner’s own citation 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). Petitioner highlighted the metrics it relies on.

The only metrics 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.* A POSITA would not understand any of these metrics to be

‘social connections,’ as claimed in the ’897 patent, *i.e.*, the in-degree and out-degree a given user has to other active users in the social network. EX2008 at ¶173. A resultant disadvantage of *Phillips* is that, without considering “social connections,” the analysis could be thrown off by common behaviors. *Id.* “For example, sending lots of emails to very few of your connections and/or receiving lots of emails from a small number of connections would significantly skew *Phillips*’ results.” *Id.*

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,” because neither the SNA techniques nor any other disclosure in *Phillips* consider social connections. EX2008 at 174.

Notably, the “social connections” claim element was added during prosecution and was explicitly relied upon by the examiner to allow this claim. EX1002 at 471, 472 (amendment) and 517-523 (notice of allowance).

***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;***

***(1) No Classification Models or Class Labels***

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 "exemplary classification models that predict 'class labels.'" Pet. at 46. However, as Dr. *Medvidovic* confirms, "*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." EX2008 at ¶175. "Neither classification models nor class labels are necessarily inherent in training an AI engine." *Id.*

The mere fact that *Smyth* seeks to predict the likelihood of exploitation of system vulnerabilities (*see Smyth* Fig. 1) does not teach or suggest how *Smyth* accomplishes this task. As Dr. *Medvidovic* notes, the "[u]se of classification models and class labels are not necessarily required and can therefore not be inferred." EX2008 at ¶176.

***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,***

***(1) No Likelihood of Exploitation As Claimed***

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). This is not correct.

First, as explained in element 1.d, *Smyth* does not teach or suggest either classification models nor class labels, and “neither is necessarily inherent in merely training an AI engine.” EX2008 at ¶178.

Second, as explained in 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. As explained by Dr. Medvidovic, “this is a binary choice: guaranteed-not-to-happen vs. guaranteed-to-happen.” EX2008 at ¶178. “That binary choice does not disclose the claimed ‘likelihood of exploitation’ of the ’897 patent, because that term requires a gradient.” *Id.* “All vulnerabilities with the class label “exploited” are treated equally by Petitioner’s interpretation of *Smyth*, and no likelihood value is assigned.” *Id.*

Dr. Medvidovic further explains this binary choice and its difference from the claimed invention:

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.

EX2008 at ¶180.

The '897 patent explains many times that the “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.” EX2008 at ¶181. It even indicates a stepwise increase. *Id.* (citing ’897 patent at 4:5–7). “The ’897 patent further discloses “ranking vulnerabilities,” which would be impossible in a binary classification model.” *Id.* (citing ’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.***

*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. *See* claim construction of “disclosure date” in Section VII.B (citing ’897 patent at 12:35–39).

Petitioner identifies *Smyth*’s disclosure of “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 include or reference the disclosure of vulnerabilities at all. Petitioner provides no support for its inference that this disclosure necessarily implies that *Smyth* discloses predicting “an actual exploitation of the respective software vulnerabilities before disclosure.”

Petitioner further cites Dr. Bambos in support. Pet. at 51 (citing Bambos at ¶¶119–123). But as Dr. Medvidovic points out, even Dr. Bambos admits that “some vulnerabilities are disclosed immediately upon discovery.” EX2008 at ¶185 (citing 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.” *Id.* “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.’” *Id.*

*Smyth* does not disclose this claim element, because it does not disclose whether the “likelihood of exploitation predicts an actual exploitation of the respective software vulnerabilities before disclosure” or not.

Notably, the “before disclosure” claim element was added during prosecution and was explicitly relied upon by the examiner to allow this claim. EX1002 at 471, 472 (amendment) and 517-523 (notice of allowance).

## **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.

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.

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.

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.

The '897 patent defines the “exploitation date” of a vulnerability as the date the exploit was first detected in the wild. *See* claim construction of “exploitation date” in Section VII.C (citing '897 patent at 9:42–43). *Smyth*'s disclosure does not refer to an exploitation date. Therefore, *Smyth* does not disclose detecting “vulnerabilities that appear before an associated exploitation date.”

Petitioner argues that this claim element is disclosed by *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. But, because *Smyth* does not reference an exploitation date anywhere, it is unclear whether the current vulnerability information 125 and online discussions 240 in *Smyth* are before or after an associated exploitation date. As Dr. Medvidovic explains, "it is possible that a vulnerability is discovered and discussed online only because of an exploitation of that vulnerability." EX2008 at ¶190.

Therefore, and as explained in element 1.f.1., 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.

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]).

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.

Petitioner references its claim 1 analysis for the claim element applying “a learned function to data associated with a software vulnerability to estimate a likelihood of exploitation of the software vulnerability.” Pet. at 58. Patent Owner therefore incorporates the corresponding response to claim 1. *See* section VIII.A.1.

Petitioner further references its analysis for 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. Patent Owner therefore incorporates the corresponding response to claim elements 1.f and 1.f.1. *See* section VIII.A.1.d and e.

As explained 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).

For these reasons, the combination of *Smyth and Phillips* does not teach or suggest claim 14 to a POSITA 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.

As explained regarding claim element 1.d, the combination of *Smyth* and *Phillips* does not disclose a “classification model.”

As explained 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.

As explained 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;*

Petitioner only references claim element 1.b. Patent Owner therefore incorporates its 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;*

Petitioner only references claim element 1.d. Patent Owner therefore incorporates its 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.*

Petitioner only references claim element 1.d. Patent Owner therefore incorporates its 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.

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.

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

A POSITA would not have been motivated to combine *Smyth* with *Phillips*. EX2008 at ¶¶207–211.

As Dr. Medvidovic explains, “*Smyth* identifies no problem with its system that *Phillips* would solve.” EX2008 at ¶207. *Smyth* and *Phillips* do not address the same problem and do not aim to achieve the same goal. *Smyth* discloses a method to predict the likelihood of vulnerability exploitation. *Id.* *Phillips* discloses identifying ‘important’ members of a given discussion forum based on metadata analysis. *Id.* Based on that difference alone, “there would have been no reason for a POSITA implementing *Smyth* to look to *Phillips*.” *Id.*

Dr. Medvidovic further explains that “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.” EX2008 at ¶207 (“Metadata, by definition, does not comprise ‘hack communications.’”). A POSITA would have gathered from *Phillips* that adding analysis of the communications content (*i.e.*, hacker communications) was complex to implement and even harder to scale. *Id.* (citing *Phillips* at 97 (analysis of communication content relies on “sophisticated sentiment analysis techniques, which are difficult to scale.”))

Petitioner further references EX1013 (“*Benjamin*”) in support. Pet. at 39. Benjamin also does not provide a solution to perform the complex task of communication content analysis. EX2008 at ¶208. “*Benjamin* also does not consider the communication content.” EX2008 at ¶208 (citing *Benjamin* at 7 (considering only Average-Message-Length, Number-Of-Replies, Number-Of-Threads-Involved, Tenure, Sum-Of-Attachments, and Total-Messages))).

To the contrary, while *Phillips* acknowledges 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.” EX2008 at ¶208. “Thus, *Phillips* teaches that incurring the cost of analyzing “hacker communications” is unnecessary because metadata analysis is sufficient to accomplish its goal.” *Id.* A POSITA would have been motivated to avoid incurring that additional computing cost when the much simpler analysis of metadata alone is sufficient. *Id.*

Petitioner further notes 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 attorney argument. It “does not consider the stark difference between communications of security professionals that openly share ideas, and hacker communications that intentionally obfuscate their

clandestine communications.” EX2008 at ¶209 (citing *MacDonald* (EX1015) (“Hacker communities are dark networks, so members explicitly protect their anonymity, which serves to counter detection from law enforcement.”)). For that reason, Dr. Medvidovic agrees with *Smyth* and *Benjamin* that “the analysis of the substance of hacker communications was a complex and non-obvious task at the relevant time.” *Id.*

Finally, “*Phillips* targets communications in Islamic Networks (*e.g.*, motivated by the 911 terrorist attacks) to identify persons of interest (*i.e.*, suspected terrorists) [and] has nothing to do with cybersecurity and vulnerable software, and is not analogous art.” EX2008 at ¶210.

Therefore, a POSITA would not have been motivated to combine *Smyth* with *Phillips*.

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

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.

Petitioner notes that *Smyth* and *Phillips* “do[] not expressly describe further processing that exploits data to remove noise and irrelevant data.” Pet. at 65.

Petitioner then identifies *Nunes* as disclosing “further processing that exploits data to remove noise and irrelevant data.” *Id.*

As Dr. Medvidovic explain, a POSITA would not have been motivated to further combine the combination *Smyth* and *Phillips* with *Nunes*.” EX2008 at ¶214. As explained in claim 1, the combination of *Smyth* and *Phillips* intentionally ignores communication content and analyzes only metadata. See claim ground 2, claim element 1.b. Petitioner confirms that “*Smyth* does not expressly discuss accessing ‘hacker communications’ in those forums” (Pet. at 38), and *Phillips* only aims to identify “important individuals” using “communication metadata alone” while ignoring the content of communications. *Phillips* at 97. Thus, Dr. Medvidovic explains, “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.” EX2008 at ¶215.

“To the contrary, *Phillips* teaches away from utilizing *Nunes* because *Phillips* intentionally avoids processing communication content.” EX2008 at ¶216 (citing *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.”)).

Therefore, the combination of *Smyth*, *Phillips*, and *Nunes* does not teach or suggest claim 4 to a POSITA.

## 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.

Petitioner only references its analysis of ground 2, claim 14. Patent therefore incorporates its response thereto. Petitioner further notes that “*Smyth* does not expressly state it uses a Random Forest tree-based model.” Pet. at 67. Petitioner then cites *Nunes* for the supplemental disclosure of a random forest tree-based model. *Id.* *Nunes* does not disclose this claim element.

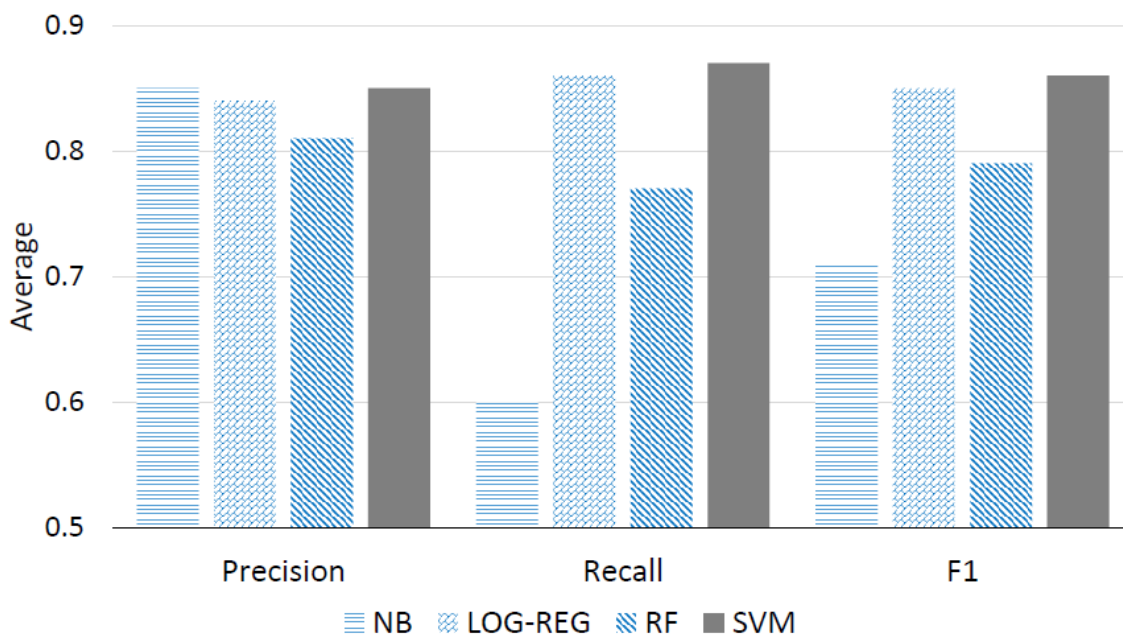
First, as explained regarding claim 4, a POSITA would not have been motivated to combine the combination *Smyth* and *Phillips* with *Nunes*, because “*Smyth* and *Phillips* do not analyze communication content, and because *Phillips* actually teaches away from that concept, in favor of analyzing solely communication metadata.” EX2008 at ¶219.

Second, while *Nunes* references the phrase “Random Forest” as a “well-known classification technique,” it does not provide any details about its use. EX2008 at ¶220 (citing *Nunes* at 2). *Nunes* provides no implementation details,

and therefore, does not teach or suggest an RF method “including multiple decision tree predictors applied in combination to classify the software vulnerability.”

EX.201 at ¶220. Petitioner does not claim this claim element as inherent in RF methods, and Dr. Medvidovic confirms that it is not. *Id.* “Other methodologies of implementing an RF method exist, for example logistic regression or Naive Bayes models.” *Id.*

Importantly, Dr. Medvidovic explains that *Nunes* actually teaches away from using RF. EX2008 at ¶221 (citing *Nunes* at 4 and Fig. 2 (“Results” section states “SVM [Support Vector Machine] 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 is supported by *Nunes*’ figure 2:



*Nunes* at Fig. 2. “In all three instances, *Nunes* shows that SVM (dark gray) significantly outperforms RF (diagonal blue stripes).” EX2008 at ¶221.

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

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.

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***

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.

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 (NLP [sic]).

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*

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.

Petitioner admits that “*Smyth* . . . does not sort and divide vulnerability data for training and testing based on time.” Pet. at 73. Petitioner then cites *Bozorgi* for 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.

Claim 8 explicitly requires a single predetermined time period from which two subsets of data are extracted. ’897 patent, claim 8. 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.” *id.* “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.” EX2008 at ¶229.

Petitioner cites to *Bozorgi*’s disclosure of “Online Exploit Detection.” Pet. at 73. Unlike 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  $t_1$ – $t_2$ .” EX2008 at ¶230. “Thus, *Bozorgi* discloses training data that always precedes testing data, in order to avoid the temporal causality problem.” *Id.* (citing *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  $t_1$ – $t_2$ ].”)). *Bozorgi*’s disclosure of a training set from an intentionally earlier time period, and a testing set from an intentionally later time period does

not disclose this claim element, where training and testing data come from the same dataset associated with the same time period. EX2008 at ¶230. Therefore, *Bozorgi* does not disclose the claimed set of training data and a potentially overlapping set of testing data from the same time period.

Petitioner’s statement that “it would also have been obvious to train and test over the same time interval using different subsets of data to evaluate the classifier’s performance” is conclusory. Pet. at 73–74 (citing *Bambos* at ¶175). Dr. *Bambos*’ cited explanation does not support the point. He only notes that any dataset must be from a time period, and therefore, all subsets thereof are from “the same time period.” *Bambos* at ¶175. Dr. *Bambos* does not address the difference between overlapping and non-overlapping training and test datasets.

## **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.

Petitioner refers to ground 2, claim element 1.c to state that “*Smyth-Phillips* teaches deriving features from various data sources 100, including features from web forums containing hacking content.” Pet. at 74. Patent Owner therefore incorporates its response to that claim element. Specifically, neither *Smyth* nor *Phillips* disclose utilizing the content of communications related to hacking.

## **IX. SECONDARY CONSIDERATIONS OF NON-OBVIOUSNESS**

Strong secondary considerations of non-obviousness support the ’897 patent. These must be evaluated in assessing obviousness, to provide real-world insight into the novelty of the invention. *Graham v. John Deere Co. of Kan. City*, 383 U.S. 1, 17–18 (1966). These important factors counter hindsight bias and are particularly important in close cases. *Id.*

### **A. Commercial Success**

The ’897 patent has been commercialized by co-inventor Dr. Shakarian in his licensed company Cyber Reconnaissance, Inc. (“CYR3CON”) starting in 2016. CYR3CON specifically implemented the patented features of the ’897 patent in its PR1ORITY platform for predictive cyber-threat intelligence. Second Declaration of Dr. Shakarian (EX2011) at ¶7. Specifically, the product incorporated some of the key systems and methods claimed in the ’897 patent for predicting software vulnerability exploitation using machine learning models trained on data from sources like hacker forums on the dark web and deep web, enabling prioritization

of patching to mitigate threats. *Id.* at ¶¶7 and 8 (“The [’897 patent] technologies were fundamental to the success of the CYR3CON PR1ORITY product”); *see also* EX2012 (CYR3CON’s PR1ORITY product roadmap) at 3, 11, 15, 16 (referencing hacker communications and multi-source data collection from the ’897 patent).

CYR3CON’s PR1ORITY platform achieved meaningful commercial success directly attributable to the features claimed in the ’897 patent and the related portfolio. EX2011 at ¶¶8, 11. The platform integrated the ’897 patent’s predictive models to assess exploitation likelihood based on hacker communications and social network metrics, providing unique proactive threat intelligence. EX2011 at ¶¶6, 7, and 11. Customers included [REDACTED]

[REDACTED]. EX2011 at ¶9. The technology was also licensed, including to [REDACTED]

[REDACTED]. *Id.* CYR3CON generated [REDACTED] before its acquisition by Cyber Security Works (now Securin Inc.) in 2022. EX2011 at ¶10.

This success is directly tied to the patented technology: the ’897 patent’s use of multi-source data integration, including dark web features, enabled high-accuracy predictions (*e.g.*, 90% true positive rate with low false positives), which

was essential for the platform's ability to identify emerging threats ahead of exploitation. EX2011 at ¶11; *see also* EX2012 at 3, 11, 15, and 16.

### **B. Long-Felt Need and Government Support**

The '897 patent addressed a long-felt need for effective prediction of cyber exploits in a rapidly evolving threat landscape, critical to U.S. national security. EX2011 at ¶6. For that reason, Dr. Shakarian's research was funded by competitive government grants totaling over \$7 million, including from DARPA, IARPA, and ONR, focused on mission-critical cybersecurity. EX2011 at ¶4. This support highlights the invention's role in solving persistent challenges in threat prediction, where prior approaches lacked the integration of dark web data and machine learning for accurate prediction.

### **C. Failure Of Others**

The prior art at the time failed to achieve the inventions of the '897 patent. Others explicitly discouraged various concepts claimed in the '897 patent. For example, *Phillips* explicitly teaches away from analyzing hacker communication content, and instead relies on metadata alone. *Phillips* at 97. Similarly, *Nunes* teaches away from using the claimed RF model. EX2008 at ¶221 (citing *Nunes* at 4 and Fig. 2). Thus, others attempted to solve the problem but failed.

#### **D. Unexpected Results**

*Phillips* and *Nunes* teach away from various aspects of the invention, confirming that the results were surprising and not predicted.

#### **E. Skepticism of Experts**

The same teaching away confirms that leading experts in the field doubted the invention.

#### **F. Praise By Others**

The fact that Dr. Shakarian's research that led to the '897 patent was repeatedly funded by competitive government grants totaling over \$7 million, including from DARPA<sup>8</sup>, IARPA<sup>9</sup>, and ONR<sup>10</sup>, supports that industry experts acknowledged the significance of the invention. EX2011 at ¶4.

#### **G. Copying By Competitors**

Skysong's suits against the CrowdStrike and Fortinet confirm that competitors have copied the invention, which suggests they recognized its value and that it was not obvious to them.

---

<sup>8</sup> Defense Advanced Research Projects Agency, U.S. Department of Defense agency that funds national security technology.

<sup>9</sup> Intelligence Advanced Research Projects Activity, R&D arm of U.S. Director of National Intelligence.

<sup>10</sup> Office of Naval Research, U.S. Navy and Marine Corps' science and technology programs.

## **X. CONCLUSION**

The Board should deny institution because CrowdStrike has failed to demonstrate a reasonable likelihood of prevailing as to any challenged claim.

Dated: November 10, 2025

Respectfully submitted,

/s/ Jay P. Kesan

Jay P. Kesan (Reg. No. 37,488)  
KEY KESAN DALLMANN PLLC  
1050 Connecticut Avenue, N.W. Suite  
500 Washington, D.C. 20036  
Email: jay.kesan@kkd-law.com  
Telephone: (202) 772-1100

Alfonso Chan  
Andrew Cochran  
Cahill Gordon & Reindel LLP  
900 16th Street, N.W., Suite 500  
Washington, D.C. 20006  
Telephone: (202) 862-8942

*Attorneys for Patent Owner*  
*SKYSONG INNOVATIONS, LLC*

## CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. §§ 42.6(e)(4) and 42.25(b), the undersigned certifies that on November 10, 2025, a complete copy of Patent Owner's Preliminary Response was filed electronically through the Patent Trial and Appeal Board's PTABE2E System and provided, via electronic service, to the Petitioner by serving the correspondence address of record.

Dated: November 10, 2025

/s/ Jay P. Kesan  
Jay P. Kesan

## **CERTIFICATE OF WORD COUNT**

Pursuant to 37 C.F.R. § 42.24, the undersigned certifies that the foregoing Patent Owner's Preliminary Response contains less than 14,000 words excluding the table of contents, table of authorities, certificate of service, certificate of word count, and appendix of exhibits. Patent Owner has relied on the word count feature of the word processing software used to create this paper in making this certification.

Dated: November 10, 2024

/s/ Jay P. Kesan  
Jay P. Kesan