Cyber Assessor: Assessment Framework to Characterize Cyber Aptitudes

A Multi-Tiered, Semi-Automated Tool for Easing Critical Organizational Challenges

Thomas J. Klemas
SimSpace Corporation
tom@simspace.com

Abstract—**Hiring, re-vectoring, and training of employees are tasks that pose an extreme challenge for Cybersecurity Officers in many organizations, and the cost of mistakes is high. As a result, some cybersecurity managers only hire personnel that they or their trusted subordinates know personally. Others are faced with insurmountable staffing deficits and must invest a nontrivial amount of subject matter experts' time and attention to aid in finding the few strong candidates from among the mass of applicants. Similar challenges complicate cross-vectoring and training of employees. In this paper, we present a semi-automated, multi-tiered platform named Cyber Assessor, which is designed to evaluate a candidate's general and specific knowledge, skills, reasoning, critical thinking, and problem solving ability. Cyber Assessor leverages advanced data analytics to achieve full mapping of performance to specialty sub-categories and thereby enable detailed understanding of an individual's areas of strength and limitations.**

Keywords – data analytics; cybersecurity; cybersecurity assessment; risk management; NIST; FFIEC; NERC.

I. INTRODUCTION

There are numerous guidance documents from governmental authorities and works in the literature that describe approaches for assessment of cybersecurity for organizations, similar to [1], [2], and [3]. However, a key component of organizational cybersecurity is related to the cybersecurity operators that defend the organization and determining their level of skill and experience is not so straightforward. All organizations, commercial and government, large and small, face the challenge of hiring, re-vectoring, and training personnel for their job specialties.

This personnel challenge seems to be compounded enormously in the field of Cybersecurity, due to the field's emergence to priority and its rapid growth. Many organizations cannot hire fast enough, and when they try to accelerate their hiring, they frequently must expend even more resources to deal with the consequences of rushed hiring decisions. Two other strongly related challenges include cross-vectoring of candidates from related fields (i.e. Information Technology) into the best matching Cybersecurity specialty, as well as allocating resources for proficiency and currency training of existing cybersecurity professionals.

Suboptimal decisions in response to these critical needs can waste scarce resources and severely escalate an organizations' cybersecurity risk. Furthermore, there are two significant, additional, and frequently overlooked impacts on resources: (1) At some point, technical staff much be engaged to interview and evaluate technical candidates, either for new hire or cross-vectoring. In many organizations, a combined effort of human resources personnel and managers is utilized to screen applications and select the candidates for interview. Subsequently, most organizations then schedule technical staff that are best suited to interview the candidate in question. Unfortunately, many weak candidates slip through the human resources/management filter, and the interview process can be time-consuming and siphons the attention and time of technical experts from their primary duties. (2) Lacking any fine-grained understanding of the workforce members' expertise, training managers typically adopt a one-size-fits-all training approach. Everyone is given all training, whether or not they need it. Frequently, this sort of untargeted, carpet-bomb-approach training is also watered down so it can fit in limited time windows and the result is both less effective and partially squandered because it includes a large fraction of employees that already understand the subject matter.

The Cyber Assessor (CA) approach and technology that we present in this paper has been designed to efficiently assess an individual, to resolve the challenges discussed in the earlier paragraphs, and to provide advanced characterization of examinee skills, capabilities, general and specialty knowledge, reasoning, critical thinking, problem solving, and persistence. Cyber Assessor achieves these core capabilities, in part, through its multi-tiered design. A powerful innovation introduced with the tiered approach of the Cyber Assessor platform is the addition of multiple dimensions, each of which measures different components an examinee's ability. The capture of multi-dimensional measurements dramatically expands the Cyber Assessor system's capability to differentiate examinees, which directly enables greater insight to improve decision making.

In a previous paper, [6], the first author presented data analytic approaches that were designed to "identify important but non-evident structural groupings, resolve community clusters, develop insights based on the evolving

 ISBN: 978-1-61208-512-8

structure and associated history, and to make sense of the raw data, the ultimate objective for Sensemaking technologies." The Cyber Assessor design team has taken a similar approach in developing advanced analytics to characterize the performance of Cyber Assessor examinees and is currently proceeding to design enhanced automation for analysis and to maximize efficiency of exam report content generation.

Cyber Assessor incorporates a relational database design that supports full mapping of every fine-grained sub-measure, whether it is a question, exercise, or complex lab problem, to every specialty category and subcategory for which it probes examinee ability or knowledge, across an arbitrary number of customer specified specialty classification systems. Current classification systems include SimSpace specialty categories, government categories, and custom customer categories. Thus, assessment reporting can be tailored to the customer's desired specialty description system and will generate a full characterization of examinee performance across every customer requested specialty categorization. This mapping and characterization capability is especially useful to understand an examinee's areas of strength and limitations, improving hiring decisions and enabling highly targeted training or retraining for maximum efficiency and improvement.

The remainder of this manuscript is arranged as described herein. Section II describes the technical details of the Cyber Assessor platform and how it achieves its objectives. Section III provides a brief description of the types of data products and a sample of result charts that are included in the evaluation report and illustrates with example examinee performance results how key insights are produced from data analytics. Section IV summarizes the key benefits of this technology and approach. Finally, the acknowledgment and reference sections complete the manuscript.

II. TECHNICAL DETAILS

The Cyber Assessor system incorporates a number of innovations that enable it to achieve a high level of examinee differentiation and deliver actionable insights to support decision making for critical actions like hiring, re-vectoring, and training. First, the CA platform consists of multiple tiers, depicted in Fig. 1, below.

Tier 1 is intended to test general knowledge, computer science, and cyber aptitude and intellectual curiosity. Tier 2 is focused on evaluating an individual's specialty knowledge in cyber subtopics, reasoning, and critical thinking. Tier 3 challenges the examinee's cyber related skills and capabilities via a series of practical lab-type problems presented in a web-based format. Finally, Tier 4 is designed to challenge and assess an individual examinee's cyber skills and capabilities on the SimSpace cyber range, which can also be used for training "near the job" or "on the job".

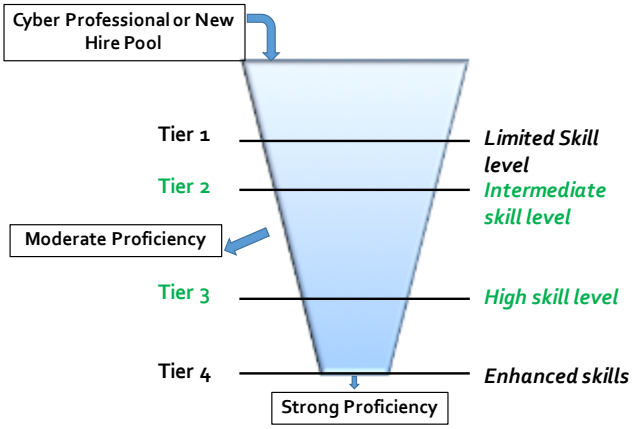

Figure 1: This figure is intended to help the reader visualize the multi-tiered structure of the Cyber Assessor platform. We highlight the Tier 2 and Tier 3 systems in green color because this paper will focus on the results and insights possible with just these 2 tiers.

Because the measures comprising each tier have key properties that are distinct from the measures of the other tiers, it is useful to think of each tier as an axis spanning a different dimension of the overall cybersecurity mastery space. The combination of the examinee performance at each tier forms a multi-dimensional score vector, as in equation 1.

$$\overrightarrow{v_{sc}} = \begin{bmatrix} v_{SC_1} \\ \vdots \\ v_{SC_5} \end{bmatrix} \quad (1)$$

It is instructive to decompose examination results to understand how the various tier 1 through 5 components contribute to describe an examinee's cybersecurity mastery. Furthermore, the CA architecture makes it possible to essentially transform coordinate systems by leveraging measure or question mappings to re-characterize examinee performance in terms of job specialty categories. This transformation starts from the individual measures or questions that comprise a particular tier scoring element. We will illustrate this transformation from tier 2 results to job specialty categories with equations 2 and 3 below. First, we show that the second element of the score arises from the tier 2 vector of measure scores which comprise scores from N_q questions in tier 2.

$$v_{SC_2} = \|\overrightarrow{v_{T2}}\| \quad (2)$$

$$\overrightarrow{v_{T2}} = \begin{bmatrix} v_{T2_1} \\ \vdots \\ v_{T2_{N_q}} \end{bmatrix} \quad (3)$$

Our goal is to transform this tier 2 measure vector into a vector of N_J job specialization category sub-scores that characterize the performance across the categories, as illustrated in equation 4.

$$\overrightarrow{v_{J2}} = \begin{bmatrix} v_{J2_1} \\ \vdots \\ v_{J2_{N_J}} \end{bmatrix} \quad (4)$$

 ISBN: 978-1-61208-512-8

To describe this characterization, we can leverage the transformation matrix, $\overline{\overline{M_{J2}}}$, which represents the mappings that comprise the relational database linkages between the measures and specialty categories, as illustrated in equation 5.

$$\overrightarrow{v_{J2}} = \overline{\overline{M_{J2}}}\,\overrightarrow{v_{T2}} \quad (5)$$

Thus, with the approach shown above, it is straightforward to transform results between multiple job characterization specialty axes. The value of designing the platform for easy transformation of this kind can be visualized in figure 2 below.

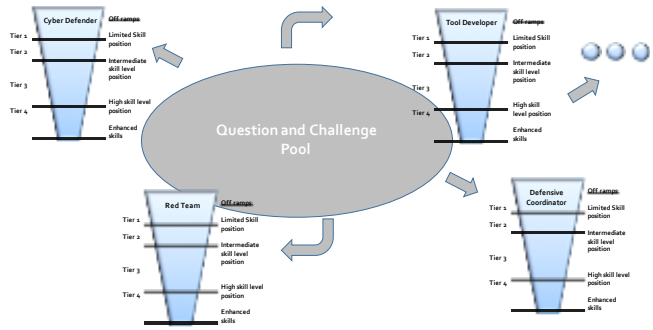

Figure 2: This figure depicts the customization enabled by flexible job specialization category mappings. A customer can identify broader job categories in which they wish to hire employees and the system will characterize performance relative to those custom job categories.

By constructing an exam with a superset combination of questions to cover every specialty category that comprises each of the broader jobs featured in Fig. 2, above, the transformations described above will be of great benefit determining a candidate's suitability for those jobs. In addition, once candidate results for multiple Tier exams are collected, it may be possible to study correlations and develop predictive tools that estimate a candidate's ability to perform on Tier 4 challenges, based simply on Tier 1 and 2 results. While this prediction approach may not be suitable for in-house candidates, in which significant resources are being invested, it can save significant resources during hiring. We also hope that these sorts of cyber assessor platform results may also enable managers to compose balanced and effective cyber teams.

The design of the measures, exercises, questions, problems also represents an area of departure from traditional knowledge retention focused approaches and an area of Cyber Assessor innovation. Although the Tier 1 and 2 exams are presented as multiple choice problems, all of Cyber Assessor's measures, exercises, questions, and problems at all Tiers were designed to probe into the examinee's fundamental skills and reasoning ability. Cyber Assessor primarily accomplishes this by posing complex challenges and evaluating the critical thinking and approach used to solve them, rather than focusing on assessing the examinee's ability to recall basic facts.

The cyber assessor analytics engine can utilize additional information about each examinee, obtained through a demographics survey, to further determine the extent of the examinees' background, level of training, and expertise and to correlate this information with exam results. The demographic survey is intended to capture relevant information pertaining to examinee backgrounds that could elucidate their performance. Towards this end, the demographic survey poses a series of questions that probe the examinee's educational background, years of interest in cybersecurity related topics years of experience in information technology, cybersecurity and various other broader areas, years of experience in a variety of narrower specialization areas, self-assessment of expertise in a variety of specialization areas, and additional accreditations or certifications that the examinee may have obtained.

The Cyber Assessor design tags each exam and each demographic survey with additional metadata that simultaneously provides their username, protects employee identity, and also link every Tier exercise and survey that the examinee completes. In this manner, it is possible to associate all of the results with the described metadata and other mappings, and these relationships are maintained in the relational database design. Furthermore, the linkages are available to the analytics engine that processes the exams, assesses performance, and characterizes the result across specialization category mappings. This capability is powerful, because frequently, it is the combination of all the available tier scoring elements with the demographic information that provides the final leap of insight as to an examinee's overall subject mastery.

The mappings illustrated in Fig. 3, below, reveals how the questions, challenges, and exercises that compose the Cyber Assessor exams and mappings encoded in the relational database are utilized by the analysis engine to characterize examinee performance across job specialty categories. The sub-result from each Cyber Assessor problem at any Tier is decomposed into that problem's contributions to each aptitude or job specialty job category, based on the relational database mappings. Thus each, examinee's results are decomposed and remapped to job specialty category sub-scores, in this example Specialty 1, Specialty 2, and each of the remaining Specialties up to Specialty N. This powerful capability enables the kind of advanced and detailed insight required to directly support leadership decision making and improve organizational cybersecurity outcomes.

III. REPORTING, EXAMPLE RESULTS, AND INSIGHTS

To evaluate the efficacy of our algorithms we conducted numerous Cyber Assessor engagements. The data collected from these engagements confirmed the effectiveness of the core capabilities that we designed into the Cyber Assessor platform. In this section, we utilize example data to share

the insights with the reader. This example data was generated artificially to avoid sharing customer data that would compromise privacy of individuals. However, the examples were carefully designed to illustrate the identical insights that we have previously achieved during the real customer engagements.

Figure 3: This figure illustrates the mapping between questions and a system for performance characterization that enables mapping to aptitude or job specialization categories. Each question is mapped to all of the categories to which it pertains.

Fig. 4, below, plots the performance distribution of examinee results from a fake organization artificially generated for this demonstration of the Cyber Assessor reporting and insight capabilities. Examinee 1 scored 100 on Tier 3 and 90 or Tier 2. Examinee 2 scored 33 on Tier 3 and 55 on Tier 2. Examinee 3 scored 59 on Tier 3 and 43 on Tier 2. Examinee 4 scored 70 on Tier 3 and 49 on Tier 2. Exam taker 5 scored 63 on Tier 3 and 33 on Tier 2. Exam taker 6 scored 100 on Tier 3 and 85 on Tier 2. Examinee 7 scored 43 on Tier 3 and 33 on Tier 2. Examinee 8 scored 22 on Tier 3 and 100 on Tier 2. Examinee 9 scored 100 on Tier 3 and 88 on Tier 2.

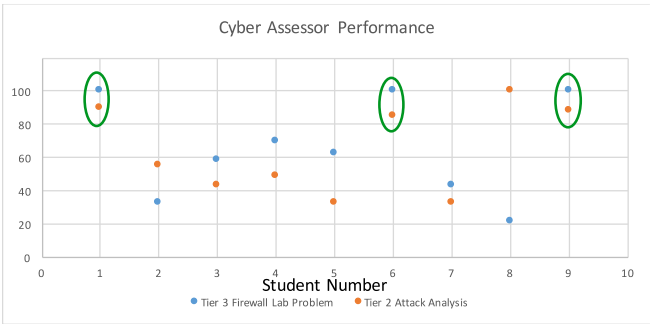

Figure 4: This charts depicts example performance data of 9 students who each took 2 exams, Tier 2 and Tier 3, is marked with green circles to indicate strong performers within the examinee distribution. These examinees, students 1, 6, and 9, performed well in both the Tier 2 knowledge and reasoning as well as the Tier 3 practical Firewall lab problem.

Since they achieved high scores in both the Tier 2 and Tier 3 Cyber Assessor platforms, this chart clearly reveals Examinees 1, 6, and 9 are high performers that demonstrated strong specialty knowledge, reasoning, skills, and problem solving, and these scores have been circled in green to highlight this insight for the reader. If these examinees were applicant candidates for hire or if these examinees were information technology specialists that were candidates for revectoring into cybersecurity, the decision maker could proceed to the next step, such as interview, with high confidence. Another potential use case: If this exercise was administered as an annual proficiency check, the leadership might consider to review these examinees as candidates for any available promotions or fast-track career programs.

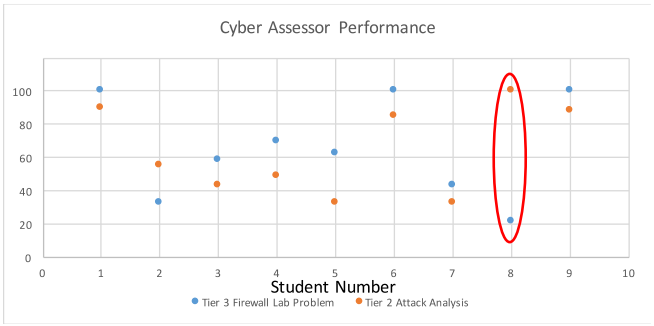

Figure 5: Sample Cyber Assessor performance data set is marked with red circle to indicate a candidate for hands-on training. Student 8 performed well on the Tier 2 knowledge and reasoning intensive exam but did not complete the practical firewall lab problem, receiving little credit.

Figure 5 represents the same results as figure 4 but is included separately to focus on an apparent anomaly that is observed in examinee 8's scores, which are circled in red to highlight this result for the reader. Examinee 8 scored a perfect 100 on the Tier 2 multiple choice knowledge and reasoning intensive exam but scored poor, only achieving 22, on the Tier 3 firewall problem solving lab exercise. This case is highly representative, not all that unusual, and is observed in customer engagements more frequently than one might initially expect. To the first glance this result seems inconsistent and it seems strange that an individual would demonstrate a high level of mastery of knowledge and reasoning and yet perform well below average in executing some of the skills that fall within his or her knowledge area.

There are several potential hypotheses that immediately come to mind when viewing the result in Fig. 3: (1) The topic of the Tier 3 exam was outside of the Examinee's expertise area. (2) The examinee was interrupted or distracted during Tier 3 exam. (3) The examinee knows a lot about his or her specialty area and has good reasoning but is very rust at actually doing things in a hands-on setting. (4) The examinee actually understood the Tier 3 problem but simply made a typo-type mistake. Several of these hypotheses, 2 and 4, can be discounted immediately by deeper dive into Cyber Assessor. Examinee exam actions can be observed in real time during the exercise, and the platform records partial results for later analysis. Either of these features are sufficient to discount that the examinee was disrupted or experienced a trivial typo-level mistake, because it is possible to observe that the examinee was

 ISBN: 978-1-61208-512-8

repeatedly attempting to solve the lab problem throughout the exercise period.

Results like those highlighted by example examinee 8 actually have really emphasized the tremendous advantage of the multi-tiered Cyber Assessor platform to differentiate examinees that results from the multi-dimensional measurement. In the cases for which this example is representative, further understanding was obtained through a demographic survey, through anecdotal evidence, from observations, and from ensuing discussions with customer leadership when reporting results from their engagement.

It seems clear that an individual who has results like examinee 8 has the basic tools (knowledge and reasoning) to perform well in their job area but might really benefit from additional training. This very well might be a valid conclusion, but the next figure, Fig. 4, reveals some additional insight that arises when the analytics taps into the demographic survey information.

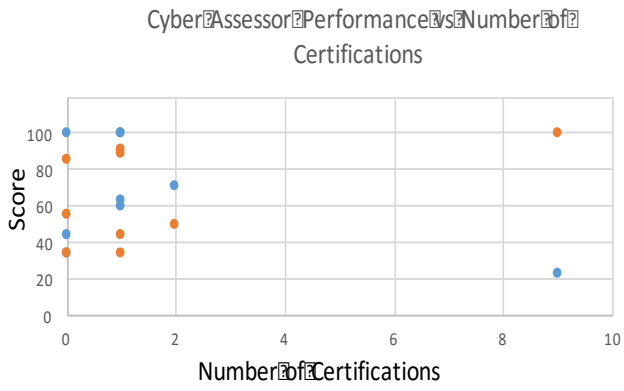

Figure 6: This chart plots example student exam performance scores versus the number of certifications that each student has achieved, captured from a demographics survey that was presented to the examinees through the Tier 2 platform. The points to the far right, corresponding to 9 certifications seems anomalous but is explained in the text of this section.

Fig. 6 plots exam performance versus number of certifications achieved by the examinee. To the far left of figure, we seem the main cluster of examinees that have accomplished between 0 and 2 certifications. The the far right, we see one outlier data pair, which corresponds to examinee 8's Tier 2 and Tier 3 scores, that indicate examinee has 9 certifications!

This is an incredible number of certifications and certainly explains examinee 8's performance on the knowledge and reasoning topics. Furthermore, it also strengthens the justification for accepting hypothesis 3. However, there are additional potential insights that a decision maker at this organization should consider. First, traditional accreditations, certifications, and exams primarily focus on examining knowledge retention and may not provide any insight into actual skill or problem solving. Skill and problem solving ability are developed through significant amounts of practice. Fortunately, CA Tier 3 is able to measure skill and problem solving. Second, given the additional information about certification, a decision would realize that employee 8 does not simply require additional training, but would benefit more from highly targeted hands-on training, individual coaching in problem solving, and perhaps additional opportunities to practice the skills associated with their specialty. We will not show an example chart that plots examinee performance versus years of experience in current job, but we ask the reader to contemplate the additional vastly different insights that would be possible if such a chart revealed employee 8 had just a few years of experience or if employee 8 had more than 15 years of experience.

Hopefully, the preceding thought experiment, following the insights from the previous charts, punctuates the asymmetric gain in value achieved by the CA platform's approach to collecting and analyzing contextual examinee background data in combination with performance data that has been designed to probe multiple dimensions of cybersecurity mastery. Finally, in figure 5, we wrap up the results section by showing how Cyber Assessor leverages the customizable specialty category mappings for each sub-measure to characterize examinee performance.

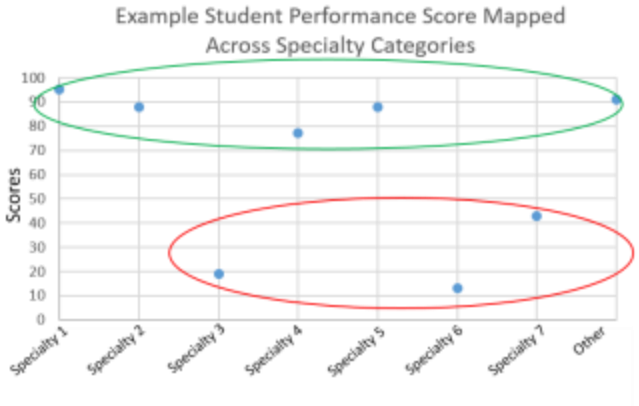

Figure 7: This chart plots an example student's performance scores versus 8 specialty categories and 1 catch-all other category to which each measure (questions, exercises, or problems) is mapped. The green circle indicates specialty categories in which the examinee performed well and might be best suited to work. The red circle indicates specialty categories in which the examinee might benefit from addition training or might not be effective to work in without additional training.

Fig. 7 maps the Tier 2 results of a particular example examinee to 7 job specialization categories, including specialties one through five and an "other" category, and the examinee scored 95, 88, 19, 77, 88, 13, 43, and 91, respectively in the measures (questions) mapped to these categories. One could imagine that the specialty descriptions might include various categories of typical cyber security operator duties such as forensics, for example. Added to the chart, a green circle was positioned to encircle scores in specialization categories where the examinee performed very well. A red circle was positioned

 ISBN: 978-1-61208-512-8

around scores in specialization categories where the examinee struggled.

Whether in support of a hiring, revectoring or training decision, the insights revealed by the chart in Fig. 5 will be of great value to a decision maker. This examinee performed quite well in the specialty 1, specialty 2, specialty 4, specialty 5, and other specialization categories but struggled with the specialty 3, specialty 6, and specialty 7, categories. Thus, if deciding about a new hire or revector candidate, the decision maker would know what area this candidate should be directed towards and which areas to avoid. The decision maker would also know exactly how to focus the training for this examinee.

IV. CONCLUSION

In this research, we have developed a multi-tiered Cyber Assessment platform that was designed to evaluate cybersecurity-pertinent skills, knowledge, and other attributes of individuals. In this paper, we presented example results that illustrated how the data analytics developed to analyze the scoring organize the performance data to maximize useful insights that support critical decision making needs of any organization. The reporting, example results, and insights section demonstrated how important insights are immediately visible in the summary charts that capture the examinee performance distribution, plot the performance against various demographic survey attribute values, and characterize individual examinee performance across arbitrary specialty categorization systems. The valuable insights, which are made possible due to the differentiation achieved by the multi-dimensional measurements that are collected by the Cyber Assessor platform and generated by its analytics and reporting subsystems, will directly support key personnel decisions. Thus, we hope that Cyber Assessor will be adopted by the organizations where it can have maximal positive impact to increase efficiency, reduce cost, and improve quality of crucial hiring, re-vectoring, and training.

ACKNOWLEDGMENT

The authors would like to thank the SimSpace corporation for presenting us with the opportunity and means to conduct this research and solve this difficult problem and the leadership for sharing insights as to challenges facing the marketplace that we addressed in this research. Finally, the authors are very grateful to Greg Gimler, John Nelson, David Carpman and David Oyer for discussions and technical contributions related to the topics in this paper.

REFERENCES

[1] FFIEC, “FFIEC Cybersecurity Assessment Tool,”, https://www.ffiec.gov/cyberassessmenttool.htm, June 2015.

[2] FFIEC, “Overview for Chief Executive officers and Boards of Directors”, https://www.ffiec.gov/cyberassessmenttool.htm, June 2015.

[3] National Initiative for Cybersecurity Careers and Studies (NICCS), “Professional Certifications”, Department of Homeland Security, https://niccs.us-cert.gov/training/professional-certifications

[4] National Institute of Standards (NIST), “Cyber Framework”, United States Department of Commerce, https://www.nist.gov/cyberframework

[5] T. Klemas and D. Rajchwald, “Evolutionary Clustering Analysis of Multiple Edge Set Networks used for Modeling Ivory Coast Mobile Phone Data and Sensemaking”, Data Analytics 2014, Third International Conference on Data Analytics.