Improving Cyber Situation Awareness by Building Trust in Analytics

Margaret Cunningham(✉) and Dalwinderjeet Kular

Forcepoint, Austin, TX, USA
{Margaret.Cunningham,DKular}@Forcepoint.com

Abstract. Analysts depend on technology to access and understand information, information that ultimately impacts their level of Cyber Situation Awareness (CyberSA). Adoption of advanced analytics, particularly those that generate risk scores or that depend on machine learning, can be impacted by a lack of trust in what the scores represent. Lack of trust in analytics can negatively impact CyberSA and efficient decision making, as analysts who do not trust outcomes from analytic models continue to search for information that confirms the analytic outcome, or continue to seek supplementary environmental information prior to making critical decisions. While human-driven investigative work is, and will remain, critical for security operations, delays in decision making, and increased efforts in information gathering, can negatively impact the efficiency of threat detection. Semi-structured interviews with analysts revealed five avenues for improving trust in analytics, including Context-Based, Case-Based, Model-Based, Ethics-Based, and Human-Centric AI Improvements.

Keywords: UEBA · CyberSA · Analytics · Risk scores · HCI

1 Background

As the cyber landscape expands through the development of more advanced malware and evasion techniques, gaps in security associated with the rapid adoption of cloud and Internet of Things (IoT) services [1], and sophisticated attacks targeting human weaknesses [2], building analytic models and technologies that can capture, consolidate, and communicate information is critical. Cybersecurity analysts depend on technology to understand the environment that they protect, which includes understanding issues such as immediate threats, malignant anomalies, vulnerabilities, and ongoing system status. This also means that how well an analyst's technology represents information directly impacts the analyst's performance in terms of their ability to assess the status of the environment and make decisions based on those assessments.

How analysts gather and process information about system status and emerging threats has changed over time, with industries transitioning away from event-based models that provide analysts with a chronological list of network events, and towards user and entity behavior analytics (UEBA) that assign risk scores to entities using analytic models. However, transitioning from event-based models to the more abstracted UEBA models of cybersecurity is not always a smooth process. One reason for this is that analysts do not always trust the analytics behind UEBA risk scores. This

W. Karwowski and T. Ahram (Eds.): IHSI 2019, AISC 903, pp. 211–216, 2019.
https://doi.org/10.1007/978-3-030-11051-2_32

lack of trust in analytics and risk scores impacts how quickly analysts can process information from their environment, as they may delay decision making to search for supplementary or confirmatory data prior to feeling comfortable enough to make decisions. This delay, and lack of trust in UEBA risk scores, can interfere with the efficiency and accuracy of threat analysis.

To identify strategies to improve trust in risk scores and UEBA, the present study provides results from a series of semi-structured interviews focused on the use of a new analytic model that relies heavily on behavioral analytic risk scores.

1.1 Event-Based and Entity-Based Analytics

For the purpose of this study, it is important to more clearly differentiate between event-based and entity-based analytics, as entity-based analytics require analysts to trust algorithms that they may be unable to fully access or understand. This is especially true for analytics that combine rule-based analytics with more complex statistical approaches.

The present study focuses on analysts' relationships with entity-based analytics. Broadly speaking, event-driven analytic workflows present users with sets of events that are most critical to review, whereas entity-driven analytic workflows present users with a set of entities that have displayed unusual behaviors and/or behaviors consistent with known risky scenarios in the form of a risk score. In the case of entity-based analytics, users are also provided with some relevant context for risk scores.

The word entity can represent more than one thing. For instance, an entity could be a user or a machine, an entity could be monitored or unmonitored, and each entity has its own meta data. The word event can also represent more than one thing, such as an email or a trade, and events can include features that make them more useful (e.g., To, Bcc).

1.2 CyberSA

An analyst's ability to respond quickly to a threat depends on CyberSA, and as attackers can do an enormous amount of damage in a short period of time, improving the speed with which analysts respond is a high priority goal. Use of the term CyberSA is based on Endsley's situation awareness (SA) framework [3] and its application to cyber security [4]. CyberSA is unique in that the first level of SA, perception of environmental information, is intrinsically linked to how the information is provided to the analyst. For analysts, this means that there is always a level of abstraction between their access to information and the ground truth cyber environment. This also means that the first challenge in developing high levels of CyberSA is ensuring that analysts trust the information they are provided.

The stages of CyberSA, illustrated in Fig. 1, includes analytics as its own block to emphasize the impact of analytics on situation awareness. The first stage of CyberSA is perception, the second is comprehension, and the third is projection. Perception includes what people gather from the environment using their senses, and can be impacted by past experiences, expectations, and attentional allocation. It is also impacted by how certain objects or information is presented. Comprehension is a

process where people build a better understanding of how information fits together through making classifications, identifying patterns, and fitting information into goal-oriented contexts. Projection is the most advanced stage of SA, and encompasses the ability to dynamically and systemically make projections about what the future state of the environment might be.

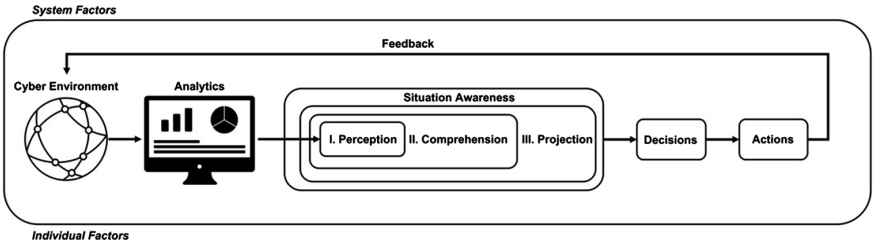

Fig. 1. CyberSA model including the layer of analytic abstraction between analysts and data from the cyber environment (figure adapted from Endsley, 1995)

2 Method

The present interview study was based on the results from a comprehensive literature review of how users and analysts interact with analytics (literature review unpublished). Findings from the literature review were used to develop a semi-structured interview template. Designing content for semi-structured interview questions by integrating findings from existing research has been shown to improve the results of qualitative interview studies [5, 6].

A total of five participants were interviewed using the semi-structured interview guidelines developed by the authors of this paper. The interviews were conducted by both authors, with each interview requiring approximately one hour. After completing the interviews, the authors independently compiled their notes, and then reviewed all findings to ensure that there were no discrepancies. Few discrepancies emerged, and those that emerged were resolved during review without issue. The consolidated results were further explored to establish thematic groupings outlined in the results section of this paper.

3 Results

Interview responses indicated a high level of agreement across five strategies that could improve trust in entity-based analytics. The strategy categories include: Context-Based, Case-Based, Model-Based, Ethics-Based, and Human-Centric AI Improvements. Table 1 provides a summary of which of the five categories each participant discussed during their interview. Details about each strategy category, including some recommendations for future improvements are also summarized below.

Table 1. Summary of strategies for improving trust in analytics by participant

	Participants					
Strategy	1	2	3	4	5	Agreement
Context-based	•	•	•	•	•	100%
Human-centric AI	•	•	•	•	•	100%
Use case-based	•	•	•			60%
Model-based	•		•		•	60%
Ethics-based	•	•		•		60%

3.1 Context-Based

Users need access to what types of data impacts risk scores in the form of context. This means that the analysts expected clarity and understanding of information regarding which policies were violated, as well as access to details regarding data movement violations. To supplement risk scores, analysts who participated in this study currently continue to reference log files to identify why a score may have increased for specific entities. Future design and development of easy-to-use interfaces that provide more context to analysts could improve efficiency.

3.2 Human-Centric AI

The analysts desired the ability to manually override any automatic actions (e.g., blocking of user behaviors, account locking) based on analytic outcomes. The analysts also wanted the freedom to create their own entity-based use cases. Providing control to the users, especially control that impacts the outcome of analytics, was highly desired and also an essential factor for building and maintaining trust with the analytic platform.

3.3 Use Case-Based

Use case-based strategies refer to providing users with information about the analytics within the context of existing use cases relevant to their area of expertise. This means that this strategy requires understanding specific business area needs. Participants noted that use cases are particularly helpful for understanding how to use UEBA, and for interpreting various features (such as timelines) presented in summary dashboards. Participants noted that case-based strategies are primarily critical during training periods, and these strategies should be integrated into all documentation and training materials.

3.4 Model-Based

Certain users or analysts desire a deeper understanding of the models that calculate risk levels. This is also associated with a desire to understand what a “high risk” level means, and how the calculations might be impacted by various data sources. For the

analysts interested in this information, providing the rationale behind and details of the model can have an immediate impact on improving trust. This information should be provided in product documentation.

3.5 Ethics

Analysts want to use data in an ethical way while simultaneously retaining the ability to do their jobs well. In light of new regulations such as the (General Data Protection Regulation GDPR), pseudonymization and anonymization of data is critical for all companies. However, some analysts share concern that pseudonymization strategies strip information out of data sources in a way that may shift the results of analytic models. As privacy concerns continue to emerge, and regulations shift, analysts as well as analytic model developers will need to build a collaborative relationship to design systems that meet both privacy and analytic requirements.

4 Conclusion

As the cyber threat landscape grows, analysts' need to rely on abstracted data processed using advanced analytics and machine learning will also grow. Analysts will also be challenged to perceive and respond to threats as quickly as possible, and begin to demand more from their tools and technologies to improve CyberSA. However, reliance on analytics will be tenuous until analysts achieve a higher degree of trust in analytic models and outcomes. While a strong relationship is possible, and likely inevitable, the community may trust more quickly when there is better collaboration between data scientists, software engineers, designers, and cyber security analysts. Meeting analytic needs through sharing expertise across disciplines will make it easier for analysts to "perceive, comprehend, and project" which will result in more efficient and potentially accurate decision-making and faster responses to threats.

The present study has limitations, most notably the small sample size, but our findings emphasize several core avenues for improving trust in advanced analytics. It is notable that one of the categories with full agreement is working towards a more Human-Centric AI strategy, which emphasizes the need for considering the human across all stages of the software development lifecycle, and the need for creating opportunities for empowering users to take control of certain aspects of analytic platforms.

As use of advanced analytics, machine learning, and artificial intelligence progresses, future research should continue to address issues such as trust, and continue to target strategies for making analysts' end-goal decisions more accurate and efficient. Analytic models that are currently under development or currently deployed may benefit from systematic research that reveals what features and capabilities analysts are really using, and what they are ignoring, and why.

References

1. Cisco Systems, Inc.: Annual cybersecurity report. Technical report (2018)
2. Forcepoint: The 2017 state of cybersecurity. Technical report (2017)
3. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Fact. **37**, 32–64 (1995)
4. Barford, P., et al.: Cyber SA: situational awareness for cyber defense. In: Jajodia, S., Liu, P., Swarup, V., Wang, C. (eds.) Advances in Information Security, vol. 46. Springer, Boston (2010)
5. Fylan, F.: Semi-structured interviewing. In: Miles, J., Gilbert, P. (eds.) A Handbook of Research Methods for Clinical & Health Psychology, pp. 65–77. Oxford University, Oxford (2005)
6. Kallio, H., Pietila, A., Johnson, M., Kangasniemi, M.: Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. J. Adv. Nurse. **72**, 2954–2965 (2016)