Considerations for Human-Machine Teaming in Cybersecurity

Steven R. Gomez, Vincent Mancuso, and Diane Staheli(✉)

MIT Lincoln Laboratory, Lexington, MA 02421, USA
{steven.gomez,vincent.mancuso,diane.staheli}@ll.mit.edu

Abstract. Understanding cybersecurity in an environment is uniquely challenging due to highly dynamic and potentially-adversarial activity. At the same time, the stakes are high for performance during these tasks: failures to reason about the environment and make decisions can let attacks go unnoticed or worsen the effects of attacks. Opportunities exist to address these challenges by more tightly integrating computer agents with human operators. In this paper, we consider implications for this integration during three stages that contribute to cyber analysts developing insights and conclusions about their environment: data organization and interaction, toolsmithing and analytic interaction, and human-centered assessment that leads to insights and conclusions. In each area, we discuss current challenges and opportunities for improved human-machine teaming. Finally, we present a roadmap of research goals for advanced human-machine teaming in cybersecurity operations.

Keywords: Cybersecurity · Cyber · HCI · Teaming · Interaction · Sensemaking · Situational awareness · Artificial intelligence

1 Introduction

With ever-increasing reliance on networked information systems, cybersecurity is a critical component of almost every military, government, and private-sector organization. As organizations deploy new technologies for their respective missions, and adversarial capabilities advance, it is clear that the goal of cybersecurity must be to *maintain* a strong defensive posture and effectively resolve incidents, rather than *achieve* some level of security and move onto the next goal. In general, this maintenance process involves mitigating vulnerable systems (including tools, people, and workflows), as well as continually observing

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.
This material is based upon work supported under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Air Force.

D. D. Schmorrow and C. M. Fidopiastis (Eds.): HCII 2019, LNAI 11580, pp. 153–168, 2019.
https://doi.org/10.1007/978-3-030-22419-6_12

and analyzing the environment for activity that might enable—or be evidence of—exploitation of both known and unknown vulnerabilities.

In the event of a security incident, analysts must determine a benign cause or understand the extent of malicious activity, identifying any adversary and their goals, their capabilities, and the intended target effects [4]. The addition of an intangible, logical cyber environment creates new complexities for analysts compared to physical security domains. For example, traditional physical domains have ecological and contextual anchors that play a role in decision-making processes, helping with validation and verification of events and courses of actions [24]; however, cyber environments lack easily observable anchors. Environments can also change rapidly and signficantly over time with few physical constraints, and change is typical even under normal conditions. To make matters worse, adversaries may take steps to hide evidence of their activities. As a result, it is difficult to design and deploy any "canary in the coal mine" for security analysts that is reliable, easy to observe and interpret, and suggests a clear follow-up response.

In fact, analysts' insights about a current security posture are primarily guided through interactions with data and mediated by computer systems, magnifying the importance of Human-Computer Interaction (HCI) challenges in this domain. Security analysts often are responsible for tasks that are very cognitively demanding: collecting, analyzing, and interpreting large, dynamic volumes of data to confirm the presence of a threat [11], while at the same time unable to prove an environment is safe with total certainty. Improved coordination between humans and machines is a promising approach for addressing these challenges but is not well understood in the cyber domain, where analyses are highly exploratory and failure to arrive at clear, justifiable conclusions is costly (e.g., failure to halt an ongoing attack).

In this paper, we explore ways to apply or enable Human-Machine Teaming (HMT), where analysts work alongside machines responsible for some duties or sub-tasks traditionally held by humans, for cyber defense. Specifically, we focus on analysis and monitoring practices in cyber defense, rather than the security of individual systems, which may highly specific to individual environments. For simplicity, we consider machine teammates in the form of software agents, or intelligent components within software applications, rather than physical devices (e.g., robots). Our goal is to understand how current challenges that analysts face could significantly benefit from using machines for Intelligence Augmentation (IA) of analysts, or as Artificial Intelligence (AI) systems that interact with analysts after performing tasks autonomously. We explore the following questions:

- What human-centered challenges exist when performing tasks with existing tools and analytics for cyber defense?
- What potential benefits can be gained from improved HMT and IA/AI during these activities?
- What are high-impact research directions that could enable these benefits?

While HMT has been studied in past systems—notably, where humans have supervisory control over unmanned vehicles [5,25,28]—it has not been studied extensively in the cyber domain. One reason could be that past HMT research assumes humans can act as supervisors who can verify and re-vector their machine teammates as needed; however, unlike vehicle control or similar applications, it is non-trivial for an analyst to supervise in a traditional sense—primarily through observation of another's activities—and verify the behavior of an analytic or agent in the cyber domain. As such, we imagine that successful HMT in cyber is as much about effective bilateral communication of complex findings as it is about task delegation or instruction.

In order to understand how HMT can improve cyber defense, and how to get there through novel HCI and security research, we contribute a set of top challenges that occur during different stages of a data-analysis pipeline for security; we outline opportunities for HMT in each stage; and we discuss implications and a research roadmap that will enable these HMT opportunities. We note that the challenges and opportunities identified are not exhaustive, but reflect key areas for improvement based on our observations of defensive cyber operations and analysis activities.

2 Model of Activities for Cybersecurity Sensemaking

In order to understand the state of security analysis challenges and where teaming can help, we consider a simple model of stages of human-initiated activities that support cyber sensemaking. Cybersecurity operations happen within a sociotechnical system, with strong interplay between humans, technology, and data. Roughly speaking, raw data must be collected and organized, then transformed by algorithms and user interfaces; then humans discover and synthesize knowledge and possible narratives that explain the data.

Teams or individuals performing these analysis activities may develop unique practices over time, but some models have been proposed to describe generally what steps are involved in cyber analysis. For example, D'Amico et al. [9] describe these security analysis activities with respect to three stages:

1. *threat detection*, where analysts collect and analyze primary sources of data;
2. *situation assessment*, where analysts bring in more data sources, and convert the analyses into actionable knowledge; and
3. *threat assessment*, where analysts look across incidents, correlating with intelligence, making predictions, and proposing mitigation strategies.

We note that while threat detection may begin with an alert generated automatically by an analytic in the environment, e.g., from an intrusion-detection system (IDS) like Bro/Zeek [1], these tasks are primarily driven by people. In some ways, these tasks mirror steps taken to operationalize data for a particular use case (here, cybersecurity): from raw data to information to knowledge, sometimes using analysis products like visualizations as inputs to later analysis stages [6].

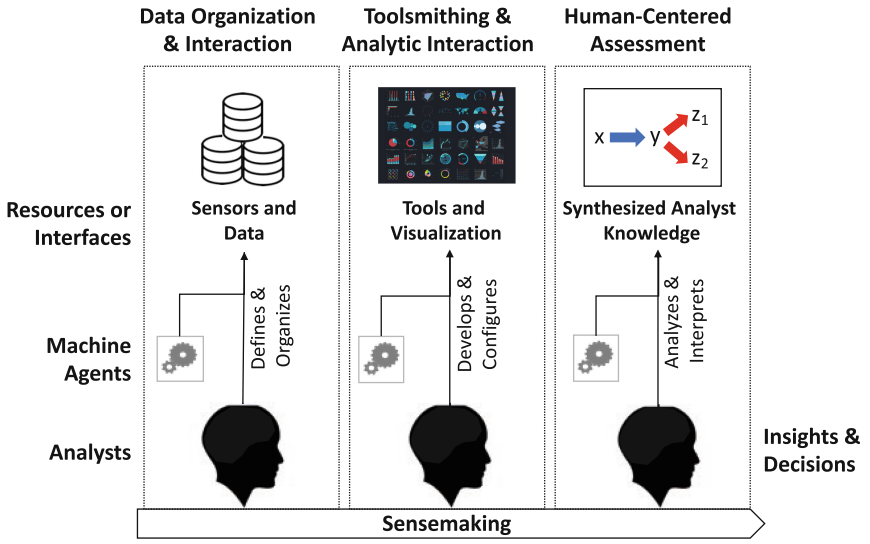

Fig. 1. Descriptive model of cyber sensemaking activities at three levels: data organization and interaction, toolsmithing and analytic interaction, and human-centered assessment.

In the remainder of this paper, we consider a model of cyber sensemaking based on the one by D'Amico et al., but generalized modestly to underscore the types of resources, technologies, and interfaces needed in each stage. Analysts sometimes must pivot across tasks and hypotheses, so have we separated out goals (which can be nested and held in parallel) from information and HCI affordances. In fact, threat detection and assessment activities are highly iterative, so analysts doing threat assessment might discover capabilities of an adversary that cause them to go back to the detection activity. As shown in Fig. 1, our model includes:

1. *data organization and interaction*, where analysts organize cyber data feeds and perform data-wrangling activities like filtering and cleaning;
2. *toolsmithing and analytic interaction*, where analysts use tools like visualizations to interpret information that has been transformed by algorithms or analytics; and
3. *human-centered assessment*, where people work with this information to construct and communicate high-level knowledge about threats or an environment.

At each stage of this model, there are critical human-centered activities that make use of machine interfaces and agents, ranging from graphical user interfaces to alerting tools that run without regular human guidance. We believe many activities can be improved beyond the current state of the art using machines that further augment analysts' performance, enable new analyses, or lighten

analysts' workloads. In the following sections, we describe existing challenges and new opportunities in each of these stages.

3 Considerations for Data Organization and Interaction

Data organization is an early stage in the analysis pipeline that is critical for downstream activities, like analysts assessing normal or abnormal conditions in the cyber environment and responding. By "organization", we refer to analysts' ability to gather and structure data in a manner that is suitable for further analysis. This process might also include interaction with data in order to clean, filter, and otherwise prepare them for analysis. While some tools may be used to interact with raw data at this stage, we distinguish those from tools developed for analysts with the goal of extracting actionable information from the data, which is discussed in Sect. 4.

Typical examples of raw data include event feeds from Security Information and Event Management (SIEM) tools and databases; records of network flows; hardware and software inventories on endhosts; and directory services for users, among others. In order to support effective sensemaking about the security of the environment, the data must capture the most relevant information about potential threats, and the data feeds themselves must be protected from compromise that could poison downstream analysis.

Understanding what data must be collected and ensuring that feeds themselves are operational (sometimes using simple analytics or monitoring tools) are important ongoing security tasks for cyber defenders. Due to the sensitive nature of data collected, another consideration is that necessary data must be available to downstream analytics and analysts, and that proper data hygiene, like archiving and protecting confidentiality, is maintained.

3.1 Challenges

Collecting High-Integrity Data. Planning for effective collection and maintaining data feeds is cognitively intensive. Along with structured information, people form mental models of operations in the environment, as well as the needs or mission of the environment, in order to plan for and understand sensors. But this model might be incomplete or become inconsistent with changes over time. Unnecessary or incorrectly-configured sensors can result in an overwhelming amount of data that is costly to manage or impacts operations—for example, by disrupting users or reducing the performance of systems like networks or endhosts. At the same time, blind spots in the network might also form and result in incomplete analyses or, worse, lead to misguided conclusions (i.e., incorrectly "clearing" a potential attack vector that remains vulnerable or exploited).

Maintaining high awareness of operations in order to address sensor issues cannot easily be solved with off-the-shelf solutions. Resources to acquire (or develop) and deploy sensors that blanket the environment—and to manage the potential deluge of data from them—can be prohibitively expensive or result in

low signal-to-noise for later analysis. Stakeholders for an environment must make choices about what to observe and when, using previous knowledge and intuition. This kind of task is important enough that it was featured in the 2016 VAST Challenge [32] and included determining which one of several data streams to enable mid-way through the exercise, given observations about the environment in the time leading up to the choice. It is not obvious how to construct automatic decision systems for these cyber choices.

Organizing Information. Cleaning and organizing raw data into useful information schemas is a critical part of the early stages of data analysis, and is both time consuming and demands advanced expertise. In some cases, this "data wrangling" can account for 80% or more of analysts' time [19] and cannot easily be outsourced, as it requires both domain knowledge and technical proficiency akin to programming [16]. Even aligning different time-series data, which might appear to be an easy task, can be challenging due to how different cyber data types are reported. For example, vulnerability reports are snapshots at a set point in time, and streaming data sources operate in real-time. Other activities require understanding the meaning and utility of data later in its lifecycle; for example, data might be removed due to resource requirements or hygiene practices, and one must decide what is safe to purge without impacting ongoing or future analyses of security incidents. Stakeholders for these data usually rely on automated approaches using simple heuristics (e.g., log rotation) due to the volume and rate at which new data are collected, even if there is some chance an old record might be needed at a later time. Furthermore, conventional ways of indexing records by time make it difficult to understand potential relationships of interest between aged-out data and preserved records.

Managing Access. Protecting confidentiality of data is challenging. Providing too much access to data can threaten security broadly, while providing too little access can prevent human defenders or analytics from observing and addressing potential security issues elsewhere in the environment. Even methods that aggregate data from multiple source might reveal sensitive information or allow it to be inferred. Current solutions to these problems usually involve both people, policy, and automated systems, where usability at the interface is a critical concern.

Ensuring that data is available only to those who are authorized and need it typically involves using access-control mechanisms that rely on curated rules that map user roles or attributes to needed resources. Methods for authoring rules that are usable by human operators have been studied in prior work [3,20], but less understood is how best to inform rule curators about access needs in the network as they evolve over time. Users requesting additional access may not understand the access-control system well enough to clearly state their needs. Others who no longer need access to resources may never proactively request removing this access if keeping it does not hinder their new objectives. Finally, in cases where access control is not enforced but sensors can assess risk or detect violations of confidentiality (e.g., identifying ongoing or completed exfiltration

from data stores), presenting actionable and timely information can be difficult: alerts can overwhelm analysts or provide too little precision to be useful.

3.2 Opportunities for HMT

Automatic Assessment of the Data Platform. Machines are particularly well-suited for workloads that involve monitoring changes in data volume or velocity, so there may be opportunities to use machines to identify where additional sensing might be desired based on statistical patterns—even if understanding the nature or intent of these changes must be determined by human teammates. There is a need for visual analytics that present well-justified recommendations for data handling and provide what-if analysis capabilities for these recommendations. Implementing changes to sensors is another area where HMT is important, because the autonomy of a machine for deploying sensors could be limited by its physical interface. For example, installing new sensing applications into a network controller for a software-defined network is currently feasible by a software agent, but installing physical proximity card readers would require human assistance or robotics.

Regular testing of sensors and the data platform in the environment, which might otherwise be tedious or difficult to repeat without human error, could be performed routinely by a machine teammate. An example might be regularly initiating events in the sensed environment, like network flows, that are expected to be detected and recorded in data storage, then verifying these records as expected. The machine could escalate alerts about unexpected behaviors to stakeholders quickly, or maintain a digest of normal test outcomes in order to avoid alerting human teams without any required action. This process parallels current approaches for automated testing and building of software tools, among others; however, determining that an outcome is normal in the environment is likely to be more involved than running unit or integration tests in controlled test environments.

Shared Representations of Mental Models and Goals. Research toward developing comprehensive sets of structured data types, tasks, and goals for cyber analysis could facilitate closer interaction between machines and humans, for whom externalizing mentals models is typically very difficult. Heer notes that shared representations let both parties "contribute to, and adaptively learn from, solutions to shared problems" [16]. Enabling the analyst to more easily verify the representations a machine is working with also builds trust. We believe these representations could also help provide domain-specific ways in which data feeds might be organized—for example, to more automatically associate incoming observations or removed data with ongoing analysis cases.

Smart Data-Access Monitoring and Control. Another opportunity exists to leverage machines for fine-grained access-control maintenance. Regularly revisiting and verifying resource needs for users or agents (e.g., through interactive

confirmation) in an environment can be tedious or error-prone for human operators if permissions are more fine-grained and change periodically, but machines could perform this ably. Furthermore, machine teammates could integrate analytics that enable these interactions to be more targeted, like cross-referencing a user's current access authorizations against their actual use based on observations in the environment. This use of HMT might free operators in charge of access control from performing maintenance tasks (e.g., or enable this practice) and providing more time for complex access-control strategies, like policy development or network management and segmentation.

4 Considerations for Toolsmithing and Analytic Interaction

The next level in our sensemaking pipeline involves people using interactive, analytic applications like visualizations to analyze the cyber data, as well as the developers who produce these tools. For some types of analysis, and for analysts with programming or scripting proficiency, the same person may assume both roles. In other cases, developers must understand enough about the available data and how analysts might use it to design effective tools.

Machines play a role here primarily as tools that transform and present data to human analysts. These tools act more as "teammates" in HMT as they perform tasks that go beyond what is precisely directed by their user. Some machines may initiate interactions with humans independently, for example by identifying patterns that would be difficult for an analyst to notice and escalating alerts for people to explore and verify. Other machines may be purely reactive, running analytic routines on data after analyst-driven interactions (e.g., visual analytics). In both cases, machines must be trusted by their teammates and communicate or display information effectively; otherwise, the added work or liability—if findings are not reliable—of using these machines threaten their long-term adoption by analysts.

4.1 Challenges

Trusting Integrity of Analytic Tools. Tools that operate with integrity in this stage transform and present information to the analyst in a way that does not cause a misleading interpretation about what is happening in the environment. Buggy implementations of tools threaten integrity, as do poor interface designs or visual encodings of data. Information displays must be legible so analysts can decode the information, interpret it, and integrate it into a larger narrative.

The ability to trust that a tool has integrity is critical, but it is often impossible or impractical to verify a tool's correctness based on its source code. Spot-checking that an analytic produces the expected output is difficult in practice (outside of testing) where ground truth is expensive to learn and might require additional tools that must be trusted themselves. As a result, analysts use tools mindful of the fact that they could be misleading. In fact, visualizations can

be misleading, unintentionally or even deliberately, by using encoding methods that subvert an analyst's ability to draw conclusions from the data [29]. In cases where the encoded data has high volume or velocity, as in streaming analytics, an analyst can also be misled if the tool minimizes or hides relevant information before it is decoded by the analyst.

Anticipating Effective Tool Designs for Humans. Tool developers must understand enough about human capabilities to build tools that are legible and interpretable. This means designing tools that work effectively in consideration of perceptual and cognitive factors that could affect an analyst's ability to decode information from the interface, which could be text-based, a visualization, or use another output modality like sound.

Some design techniques have been developing as a work-around for perceptual limitations. For example, in many use cases of visual analytics, the number of events in a dataset exceeds the number of pixels available to encode the data. (In cybersecurity, sensors may collect data over months or years that could be relevant to a single incident, such as a sophisticated network intrusion.) Existing visualization approaches for handling this volume on-screen include Focus+Context techniques [7,21], which combine an overview containing broad context with user-directed exploration for fine-grained information, and interactive views that support Shneiderman's mantra of "overview first, zoom and filter, then details-on-demand" [30]. In cases where a subset of data is presented, it is also important that a tool is clear about what data are excluded.

Model-driven visualization design—by modeling tasks, humans, and data, and making decisions based on simulations of user performance—is a compelling idea to account for human factors during toolsmithing. However, there are few time-tested human performance models that are mature enough to guide design decisions. Principles like Fitts' Law [22] and design-evaluation tools that utilize cognitive modeling (e.g., CogTool [18]) can guide simple UI design decisions (e.g., optimizing mark size or position), but generally modeling visualization effectiveness is notoriously difficult, especially for exploratory data analysis (EDA) tasks [13]. As a result, design practices often rely on gaining intuition about the application area (i.e., cybersecurity) and iterating with expert users to refine tools, which can be time consuming if done with proper rigor.

Designing for Partial Analysis and Knowledge Transfer. During security operations, analysts often must hand off findings to another person (e.g., during shift changes) for continued exploration and as context for future events. One obstacle is the difficulty of communicating one's mental model of a situation or environment. Tools are needed that go beyond exploratory data analysis (EDA) and help analysts compose narratives that include hypotheses and findings, estimates of uncertainty, and an accounting of what data was analyzed or not. Maintaining rich histories of these analysis records could pose technical challenges. Partial analysis products may need to be compressed or updated when later information is available.

4.2 Opportunities for HMT

Provenance Tracking in Visual Analytics. We believe an important step for HMT is designing mechanisms for analytic provenance in order to support trust in machine teammates and their products. Analytic provenance is traditionally about understanding through interactions with analysis tools how humans arrive at insights [26]. This is very important in order for analysts to establish trust in others findings. Machine teammates must also endeavor to provide evidence that their analyses have been executed in correct and justified ways, in order to get buy-in from human teammates or supervisors about conclusions or recommendations. Ways of building provenance tracking into tools and algorithms to use representations of provenance (i.e., usually large graphs) are rich areas to explore, especially since complex analytics might involve machine learning or other approaches that are difficult to explain on a step-by-step basis. Make machine learning "explainable" to the analysts who depend on them in HMT—not just model developers and toolsmiths, as recent work has focused on (e.g., [35])—is an important future goal. Part of communicating provenance also includes effectively describing uncertainty in the analysis [36], which is an ongoing research area in information visualization.

Living Notebook and Narrative Visualization. Building on provenance, there is an opportunity to use visual analytics that capture both human inputs and findings by machine analytics into a narrative that can adapt over time. This would support ongoing analyses and knowledge transfer between teammates with less context about previous events. New tools like "living notebooks" [8] that evolve over time are for potential method handling streaming data, which must be quickly integrated into existing cases or analyses. Annotations and feedback provided by humans could be used to refine the narrative, while machines could learn from this feedback to better handle future data.

User-Performance Modeling for Cyber Tool Design. Modeling how well a visualization or other tool might support an analyst's cyber task could supplement existing ways for designing effective tools in this domain, which include design studies (see [23] for examples) that are valuable but expensive to perform. As we mentioned earlier, modeling tools can be used to get fast, quantitative predictions on performance indicators like task speed. However, effectively modeling cyber tasks requires more research because they tend to encompass both routine interactions (like pulling up and searching logs) that are straight-forward to model, as well as exploratory or less-structured brainstorming tasks. Reusable modules that instrument user interfaces, both for downstream model fitting and other performance monitoring, would be useful for visual analytics and other HMT interfaces.

5 Considerations for Human-Centered Assessment

While advanced technologies may be responsible for collecting, reducing, and processing data during initial analyses, human analysts are the primary drivers

of cybersecurity understanding and sensemaking today. They interact with analytics and other analysts to produce information and knowledge for the purpose of situational awareness and decision making in the organization. To be effective at this stage, analysts must be able to create linkages between information about the network, the world, and their team [14]. This information must be correlated with external information on emergent threats and threat actors, as well as known attack signatures. Finally, analysts must also be able to fuse and share this information with their local and broad organizational teams to create a holistic picture of security across the organization [31].

5.1 Challenges

Multi-source Information Fusion. At the individual level, analysts must correlate information between multiple sources to produce knowledge, and communicate this knowledge to their superiors. As information and knowledge are passed up through the organizational hierarchy, findings from multiple analysts must be translated from discoveries to insights, and eventually into a broader picture. This process of information fusion lets analysts achieve improved accuracy and understanding, compared to looking at an individual source of information. The fusion occurs over five levels: data assessment, object assessment, situation assessment, impact assessment, and process refinement [34]. At the analyst level, this is often discussed as "hard" information fusion, in which the focus of data is from hard sensors collecting objective information. As information is moved up the organization, the fusion moves to "hard/soft" where the hard information is fused with subjective information, which might be more uncertain, inaccurate, or subjective [15]. At all levels, information fusion is a cognitively-demanding task that requires memory, merging and conflict resolution, and de-confliction to ensure that final conclusions are accurate and actionable.

Information Sharing Across Organizational Structure. As analysts process information and reveal incidents or other status indicators, they are responsible for communicating this information up the chain for the purpose of awareness and decision making. Before doing so, the analyst must make a judgement call of whether or not a piece of information should be shared. If the analyst shares too much, she may cause information overload to her superiors; on the other hand, if she does not share enough, this degrades the situational awareness of the organization. The decision can be stressful or cognitively taxing. Research has shown that humans are more likely to share commonly-known information, while high-value, unique information they possess is not communicated [17,33]. At each level of the organizational hierarchy, information and knowledge is further distilled, fused with other information, and summarized. Where an analyst may be responsible for assessing an individual security event, his supervisor will have to understand the interdependencies across multiple events, look for patterns, and understand how to allocate resources.

Performance Measurement. Individual and organizational behavior requires monitoring and self-regulation of their actions, in order to adjust for emergent threats or to improve their overall performance. Regulation of behavior based on performance is a meta-cognitive task, in which an analyst must monitor her own cognitive behavior for task-specific knowledge, her understanding of that knowledge, and her affective responses to the activity [10]. Understanding one's performance and competencies in a particular area is a critical element in enabling trust and team dynamics [2], and is useful in assessing performance of individuals and teams. Additionally, without such information, supervisors cannot correctly allocate resources or balance the workload across their teams, which could help increase team performance [12].

5.2 Opportunities for HMT

Intelligent Information and Context Fusion. Information fusion requires pulling and aggregating findings from multiple sources, both hard (e.g., data) and soft (e.g., analyst reports) to form a summative understanding of the broader organizational picture. With tools that help create linkages between the information an analyst receives, the sources of the data and their trustworthiness and constraints, analysts can build more context around the information they are provided. Current workflows for building context like this can be ad hoc and use many tools. Unified interfaces that help synthesize and share knowledge and hypotheses between team members could lead to more systematic or streamlined analyses.

We previously discussed the difficulty in sharing information across analysts and the organizational hierarchy. Without context (like threat or analysis priorities that are communicated top-down by decision makers), it is difficult to know what information to share upwards; but without more information, it can be difficult to refine or understand some contexts. Natural language processing could help this issue by helping making it easier for analysts to construct context. As mentioned earlier, tracking sources of data and analytic provenance can help a person receiving synthesized information to learn how it was generated. This could help reduce potential data overload, allowing analysts to better understand and communicate their needs, and ensuring that information that needs to be communicated and shared.

Performance Monitoring Capabilities. Current research in neuroergonomics and physio-behavioral monitoring is making significant advances in developing metrics of fatigue, stress, and other state-based metrics that are linked to human interaction and performance while using technology [27]. Machines can use this information to augment a supervisor's intelligence and assist, or automate, tasks like load scheduling, resource allocation, and workload balancing across the team. Additionally, performance-measurement outcomes may be used to communicate information about analysts objectively up the chain; this can aid a supervisor in composing teams and allocating training. Similarly, there is potential to use HMT in situations where machines passively observe individual differences and strengths, and provide suggestions to leadership for how best to deploy teams.

Table 1. Research roadmap for improved HMT in cybersecurity

When	Research goal	D	T	HCA
Near	Methods for guiding data collection; curation assistance for those with little or no developer experience	✓	.	.
	Task modeling and representation for cyber defense operations	✓	✓	.
	Organizational Knowledge Management capabilities for intelligent information sharing	.	.	✓
	Explainable machine learning (ML) for analytic developers	.	✓	.
Mid	Improved visualization and analytics that provide distilled narratives of multi-dimensional change over time	.	✓	✓
	Accurate models for human performance in cyber defense	.	✓	✓
	Human cognitive and affective state detection	.	.	✓
	Tools for tracking and communicating analytic provenance	.	✓	.
	Natural language understanding for precise analysis tasks and wrangling data	✓	.	✓
	Support for externalizing and sharing mental models of an environment and analysis goals	✓	.	✓
Far	Explainable ML available for analysts using ML-based analytics	.	✓	✓
	Human cognitive augmentation for performance improvement	.	.	✓
	Tools for operations that adapt to individual needs and team composition	.	✓	.

6 Roadmap

The purpose of this paper was to present observations about the current state of cyber sensemaking activities, their associated challenges, and suggest opportunities for HMT in this domain. We believe the security and HCI communities can advance toward these opportunities by pursuing a research agenda at that intersection. In Table 1, we summarize some objectives in line with this agenda. This table is not meant to be complete or the product of a rigorous research-space analysis; rather, it describes some milestones related to the challenges in this paper that we think are achievable within the near (2–5 years), mid (5–10 years), and far (10–20 years) time frames. Each direction corresponds to one or more of the sensemaking stages discussed earlier: data organization and interaction (D), toolsmithing and analytic interaction (T), and human-centered assessment (HCA).

Future work in this area should not simply focus on the development of novel tools and technologies; instead we urge researchers to take a problem-based approach to addressing challenges in cyber sensemaking and analysis. Our intuition is that this will involve more closely-integrated HMT, so we can allow humans to

focus on tasks that leverage their strengths and improve their decision making. New capabilities can help support provenance, correlation, and communication across the different layers of sensemaking—enabling effective and rapid pivoting from each phase and supporting the analysis missions for which security analysts are responsible: threat detection, situation assessment and threat assessment. As more of this roadmap is achieved, it is critical for researchers to maintain awareness of existing and emerging challenges, ensuring that we leverage technologies like AI/IA and HMT in an effective and strategic manner.

7 Conclusion

In this paper, we considered current challenges involved in human-centered aspects of cybersecurity operations, focusing primarily on difficulties in analyzing and communicating findings about complex cyber environments. Many of these challenges result from information management and sensemaking of highly dynamic, multi-dimensional data. These activities traditionally have been driven by humans in the cybersecurity domain, where verifiably-complete understanding of an environment or incident is difficult or impossible to achieve; as such, it is critical to have clear and justifiable partial findings, which is beyond existing capabilities of autonomous intelligent agents. Other challenges related to human factors arise due to the fast-changing and cognitively-demanding work of security analysts.

To address the challenges, we identified opportunities for improved interactions and teaming between security analysts and machines. These opportunities exist in each of three stages of a cybersecurity-analysis pipeline model, including: (1) data organization and interaction, (2) toolsmithing and analytic interaction and (3) human-centered assessment at the level of individuals up through groups and higher-level stakeholders in an organization. Many of these opportunities must be enabled by new research directions in the security and HCI fields. Based on this, we outlined several priorities for researchers.

References

1. The Zeek Network Security Monitor. https://www.zeek.org/
2. Austin, J.R.: Transactive memory in organizational groups: the effects of content, consensus, specialization, and accuracy on group performance. J. Appl. Psychol. **88**(5), 866–878 (2003)
3. Beckerle, M., Martucci, L.A.: Formal definitions for usable access control rule sets from goals to metrics. In: Proceedings of the Ninth Symposium on Usable Privacy and Security - SOUPS 2013, p. 1. ACM Press, Newcastle (2013). http://dl.acm.org/citation.cfm?doid=2501604.2501606
4. Caltagirone, S., Pendergast, A., Betz, C.: The diamond model of intrusion analysis. Technical report, Center For Cyber Intelligence Analysis And Threat Research Hanover MD, July 2013. https://apps.dtic.mil/docs/citations/ADA586960
5. Chen, J.Y.C., Barnes, M.J.: Human-agent teaming for multirobot control: a review of human factors issues. IEEE Trans. Hum.-Mach. Syst. **44**(1), 13–29 (2014)

6. Chen, M., et al.: Data, information, and knowledge in visualization. IEEE Comput. Graph. Appl. **29**(1), 12–19 (2009)
7. Cockburn, A., Karlson, A., Bederson, B.B.: A review of overview+detail, zooming, and focus+context interfaces. ACM Comput. Surv. **41**(1), 2:1–2:31 (2009). https://doi.org/10.1145/1456650.1456652
8. Cook, K.A., Burtner, E.R., Kritzstein, B.P., Brisbois, B.R., Mitson, A.E.: Streaming visual analytics workshop report. Technical report PNNL-25266, 1417447, March 2016. http://www.osti.gov/servlets/purl/1417447/
9. D'Amico, A., Whitley, K., Tesone, D., O'Brien, B., Roth, E.: Achieving cyber defense situational awareness: a cognitive task analysis of information assurance analysts. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting 49(3), 229–233, September 2005. https://doi.org/10.1177/154193120504900304
10. Efklides, A.: Metacognition and affect: what can metacognitive experiences tell us about the learning process? Educ. Res. Rev. **1**(1), 3–14 (2006). http://www.sciencedirect.com/science/article/pii/S1747938X06000029
11. Fink, G.A., North, C.L., Endert, A., Rose, S.: Visualizing cyber security: usable workspaces. In: 2009 6th International Workshop on Visualization for Cyber Security, pp. 45–56, October 2009
12. Funke, G.J., Knott, B.A., Salas, E., Pavlas, D., Strang, A.J.: Conceptualization and measurement of team workload: a critical need. Hum. Factors **54**(1), 36–51 (2012). https://doi.org/10.1177/0018720811427901
13. Gomez, S., Laidlaw, D.: Modeling task performance for a crowd of users from interaction histories, pp. 2465–2468. ACM, May 2012. http://dl.acm.org/citation.cfm?id=2207676.2208412
14. Gutzwiller, R.S., Hunt, S.M., Lange, D.S.: A task analysis toward characterizing cyber-cognitive situation awareness (CCSA) in cyber defense analysts. In: 2016 IEEE International Multi-disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), pp. 14–20, March 2016
15. Hall, D.L., McNeese, M., Llinas, J., Mullen, T.: A framework for dynamic hard/soft fusion. In: 2008 11th International Conference on Information Fusion, pp. 1–8, June 2008
16. Heer, J.: Agency plus automation: designing artificial intelligence into interactive systems. Proc. Natl. Acad. Sci. **116**, 1844–1850 (2019). https://www.pnas.org/content/early/2019/01/29/1807184115
17. Jefferson, T., Ferzandi, L., McNeese, M.: Impact of hidden profiles on distributed cognition in spatially distributed decision-making teams. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 48, no. 3, pp. 649–652, September 2004. https://doi.org/10.1177/154193120404800380
18. John, B.E., Prevas, K., Salvucci, D.D., Koedinger, K.: Predictive human performance modeling made easy. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2004, pp. 455–462. ACM, New York (2004). https://doi.org/10.1145/985692.985750, Event-Place: Vienna, Austria
19. Kandel, S., Paepcke, A., Hellerstein, J.M., Heer, J.: Enterprise data analysis and visualization: an interview study. IEEE Trans. Vis. Comput. Graph. **18**(12), 2917–2926 (2012)
20. Krishnan, V., Tripunitara, M.V., Chik, K., Bergstrom, T.: Relating declarative semantics and usability in access control. In: Proceedings of the Eighth Symposium on Usable Privacy and Security - SOUPS 2012, p. 1. ACM Press, Washington, D.C. (2012). http://dl.acm.org/citation.cfm?doid=2335356.2335375

21. Lamping, J., Rao, R., Pirolli, P.: A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1995, pp. 401–408. ACM Press/Addison-Wesley Publishing Co., New York (1995). https://doi.org/10.1145/223904.223956
22. MacKenzie, I.S., Buxton, W.: Extending Fitts' law to two-dimensional tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1992, pp. 219–226. ACM, New York (1992). https://doi.org/10.1145/142750.142794, Event-Place: Monterey, California, USA
23. Mckenna, S., Staheli, D., Meyer, M.: Unlocking user-centered design methods for building cyber security visualizations. In: 2015 IEEE Symposium on Visualization for Cyber Security (VizSec), pp. 1–8, October 2015
24. McNeese, M.D.: How video informs cognitive systems engineering: making experience count. Cogn., Technol. Work. **6**(3), 186–196 (2004). https://doi.org/10.1007/s10111-004-0160-4
25. Mouloua, M., Gilson, R., Kring, J., Hancock, P.: Workload, situation awareness, and teaming issues for UAV/UCAV operations. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 45, no. 2, pp. 162–165, October 2001. https://doi.org/10.1177/154193120104500235
26. North, C., et al.: Analytic provenance: process+interaction+insight. In: CHI 2011 Extended Abstracts on Human Factors in Computing Systems, CHI EA 2011, pp. 33–36. ACM, New York (2011). https://doi.org/10.1145/1979742.1979570
27. Parasuraman, R.: Neuroergonomics: research and practice. Theor. Issues Ergon. Sci. **4**(1–2), 5–20 (2003)
28. Parasuraman, R., Barnes, M., Cosenzo, K., Mulgund, S.: Adaptive automation for human-robot teaming in future command and control systems. Technical report, Army Research Lab, Human Research and Engineering Directorate, January 2007. https://apps.dtic.mil/docs/citations/ADA503770
29. Rogowitz, B.E., Treinish, L.A., Bryson, S.: How not to lie with visualization. Comput. Phys. **10**(3), 268–273 (1996). https://doi.org/10.1063/1.4822401
30. Shneiderman, B.: The eyes have it: a task by data type taxonomy for information visualizations. In: Proceedings of the 1996 IEEE Symposium on Visual Languages, VL 1996, p. 336. IEEE Computer Society, Washington, DC (1996). http://dl.acm.org/citation.cfm?id=832277.834354
31. Staheli, D., et al.: Collaborative data analysis and discovery for cyber security. In: Symposium on Usable Privacy and Security (SOUPS) Workshop on Security Information Workers (2016)
32. Staheli, D., et al.: VAST challenge 2016: streaming visual analytics. Technical report, MIT Lincoln Laboratory Lexington United States, October 2016. https://apps.dtic.mil/docs/citations/AD1033423
33. Stasser, G., Stewart, D.: Discovery of hidden profiles by decision-making groups: solving a problem versus making a judgment. J. Pers. Soc. Psychol. **63**(3), 426–434 (1992)
34. Steinberg, A.N., Bowman, C.L.: Revisions to the JDL data fusion model. In: Handbook of Multisensor Data Fusion, pp. 65–88. CRC Press (2008)
35. Wongsuphasawat, K., et al.: Visualizing dataflow graphs of deep learning models in TensorFlow. IEEE Trans. Vis. Comput. Graph. **24**(1), 1–12 (2018)
36. Xu, K., Attfield, S., Jankun-Kelly, T.J., Wheat, A., Nguyen, P.H., Selvaraj, N.: Analytic provenance for sensemaking: a research agenda. IEEE Comput. Graph. Appl. **35**(3), 56–64 (2015)