ConstellationBuilder: A High-Level Situational Awareness and Team Assembly Interface for Cybersecurity Events

VAST 2020 Mini-Challenge 3: Award for Effectively Transforming Task Decomposition into Conceptual Design

Chen Guo, Jieqiong Zhao, Lu Ding, Tianyi Zhang, Weiyue Deng, Prince Owusu Attah, Xiaolei Guo, Xuan Thao Nguyen, Yunran Ju, Zhenyu Cheryl Qian, and Yingjie Victor Chen

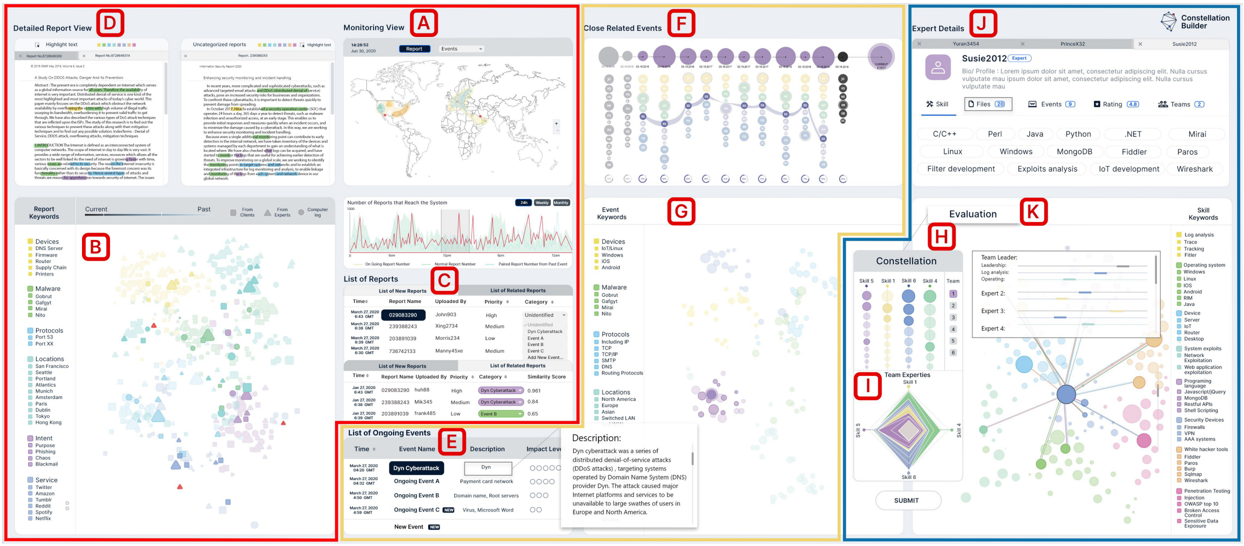

Fig. 1. ConstellationBuilder system overview: a report view (left), an event view (middle), and a team constellation view (right).

Abstract—ConstellationBuilder is an interactive visualization system to support high-level situation awareness of cybersecurity events and fast assembly of response teams. The system adopts a systematic and smooth workflow to integrate a report view, an event view, and a team constellation view into an innovative combination. It allows visual analytics experts and non-experts in machine learning to quickly and effectively analyze cybersecurity reports and events. The flexibility of the system makes it easy to extend and integrate with any other monitoring and investigation tasks in the future.

Index Terms—Situational awareness, visual analytics, cybersecurity event, expertise finding, computer-supported cooperative work

1 Introduction

A challenging task for cybersecurity analytics is to maintain a high-level network security situational awareness and quickly assemble a team of professionals with complementary skills to resolve threats. ConstellationBuilder (Fig. 1) is an interactive system we designed to support the given challenges through combining visual analytics and machine learning. Multiple visualizations are linked and varying in levels of details which enables the cybersecurity experts to efficiently investigate a new event, match it to ongoing or past cases, and assemble a team with the appropriate expertise to investigate the event. ConstellationBuilder is capable of handling a vast array of tasks and details the work necessary to support situational awareness in the event-based context and fit expertise recommendation to different cyberattack scenarios.

- *Chen Guo, Xuan Thao Nguyen is with James Madison University. E-mail: {guo4cx@jmu, nguyenxx@dukes}.edu;*
- *Jieqiong Zhao, Lu Ding, Weiyue Deng, Prince Owusu Attah, Xiaolei Guo, Yunran Ju, Zhenyu Cheryl Qian, Yingjie Victor Chen is with Purdue University. E-mail:{zhao413, ding241, deng161, powusuat, guo579, ju27, qianz, victorchen}@purdue.edu;*
- *Tianyi Zhang is with Nanjing University of Aeronautics and Astronautics. E-mail: zhangtianyi@nuaa.edu.cn.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

2 System Architecture

ConstellationBuilder's architecture is shown in Fig. 2, which aims to build the connection between cybersecurity events and experts who can resolve them. If similar cyber attacks happen in the future, ConstellationBuilder can rapidly recommend a team of white hat members. By measuring the similarity between past cybersecurity reports, the system identifies records of similar attacks and distils similar cyber events (Fig. 2 left). Latent Dirichlet allocation (LDA) [1] is used to suggest the most relevant events. ConstellationBuilder automatically identifies expertise information from incorporating documents from different sources including reports, publications, news, resumes, web pages, etc, and further leverages experts' social networks in historical events to recommend the most suitable team/constellation to respond to each event (Fig. 2 right). We call a team "constellation" because they shine like stars on solving cyber problems. By gradually refine the connection between cyberattacks and white hat members, our system can recommend teams suitable for different events. Additionally, ConstellationBuilder can assist decision makers in management roles to take more informed decisions.

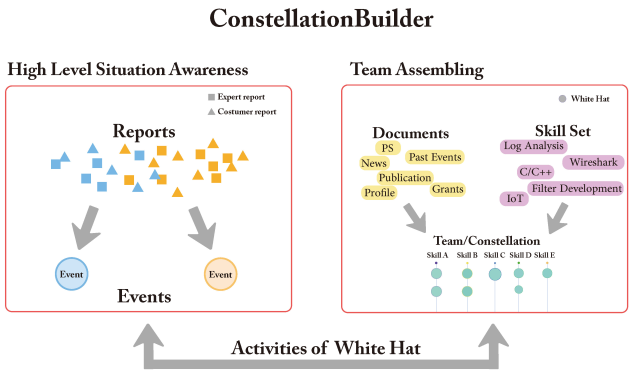

Fig. 2. The overall framework of ConstellationBuilder.

3 User Interface

ConstellationBuilder is comprised of three core displays: a report view, an event view, and a team constellation view (Fig. 1). All the views are visually linked to each other by highlighting instances of the same object in one view.

3.1 Report view

In order to uncover modern cyberattacks and mitigate their effects, it is essential to analyze the characteristic of previous similar attacks, and transform the meaningful information into actionable strategies. The report view provides insights into the current situation by monitoring the live attack reports and automatically clustering them based on the TF-IDF cosine similarity for all the documents. The report view contains four subordinate views for two purposes. A monitoring view Fig. 1A) is designed for real-time monitoring of temporal and geographical patterns. A list of reports view (Fig. 1C), a detailed report view (Fig. 1D), and a report cluster view (Fig. 1B) support the machine human collaborative identification of events.

When numerous reports, for example, experts' reports, log files, and customers' complaints, flash into the system, the system characterizes text documents by keywords/entities and keywords/entities combinations, and project them in the report cluster view (Fig. 1B) that put similar reports close together as clusters. Specifically, different shaped icons represent different types of reports; colors are used to differentiate the report clusters; the color saturation indicates the decay of time.

By integrating the report cluster view and monitoring view, the analyst can tell the difference between characteristics of reports from new events, past events, and predicted ones. ConstellationBuilder automatically classifies and ranks the ongoing reports into different events. The list of related reports view displays all the unidentified new reports and their severity statuses. Ongoing reports that are assigned with events are ordered by their similarity scores (Fig. 1C). If none of the events match this new report, the analyst can create a new event to classify the report. Additionally, in the detailed report view, reports can be investigated in more detail by changing, adding, deleting, or highlighting individual keywords, as well as assigning or changing their colors. This makes the detail report view an important and versatile tool to give more context to the current investigations.

3.2 Event view

The event view consists of a list of ongoing events (Fig. 1E), an event cluster view (Fig. 1G), and a view showing close related events (Fig. 1F). In the list of ongoing events view, the events that need to be investigated are prioritized based on severity levels. In the event cluster view, events are clustered based on their similarity and the size of the circle decodes the level of severity. Two cyber events are considered more similar if they belong to a similar event type in the cyber-incident taxonomy, happened close together in time, and contain similar keywords.

Once an ongoing event is selected, the similar past events are ordered in the close related events view (Fig. 1F) by the degree of association. Right sides are the most related events. Events colored in black indicate other ongoing events and their assigned white hat members who cannot be reallocated to the selected events. Meanwhile, the experts who handed particulars events are listed below the past events. The selected event and all the relevant keywords are also highlighted in the event cluster view. The close related event view highlights the connection between events and contributed white hat members.

3.3 Team Constellation View

By matching the keywords from the event cluster view and the skills required to resolve these events, the system automatically recommends a list of potential team members based on their previous performance, experience in solving similar past events, and their resumes, publications, etc. In the expert cluster view (Fig. 1K), the recommended white hat members are grouped based on the similarity of their skill sets. The size of each node represents the strength of overall skills. Interactions between white hat members are depicted as links between the nodes. We use different colors to show different skills and the color of each node represents the most dominant skill of a member. The skill colors are consistent with those in the event cluster view. Based on the coloring of white hat members, analysts can readily tell how many skills are required for certain events. Analysts can click on each node to examine the detailed information of each white hat member (Fig. 1J).

ConstellationBuilder views each possible team as a constellation. It recommends several constellations based on members' skill sets and experiences that can fulfill the needs of solving the cyber event as well as the strengths of social ties. The team constellation view Fig. 1H) ranks all the possible teams/constellations and their candidates. Members who have served as team leaders are highlighted in the team constellation view. Analysts can also pick members from different clusters to compose a team with complementary skills that are recommended by a systematic review of relevant events and their outcomes. For future team-building purposes, new white hat members with strong skills can be added or removed for training purposes. If no one can be picked, ConstellationBuilder will recommend one leadership to assemble a team. When a constellation is selected, the rose chart (Fig. 1I) (inspired by MetricsVis [2]) at the bottom stacks all the experts in the constellation together shows the combination of skills for a team, where each member's expertise shows as a ribbon.

After a cyber-event is resolved, each expert's performance is quantitively measured based on the number of solved exploits, average time, false-positive rate, ranking, and evaluation. The expertise levels of white hat members for this completed event are automatically recorded in the system, and the post-hoc performance analysis can assist better assignment of team members in future events.

4 Conclusion

Our proposed design effectively supports the situational awareness task and the team-building task for cybersecurity experts. The report view, event view, and team constellation view are cohesively linked together to enable a systematic workflow of identifying, classifying, investigating, and resolving cybersecurity events, even easier for a non-expert in machine learning to visually analyze them. In particular, the system architecture is flexible enough to extend to any other organizational environments and domain-expert prescribed functionalities. ConstellationBuilder can be tailored to support traditional event investigation and expertise recommendation by incorporating appropriate classification and identification modules.

References

[1] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, Jan 2003.

[2] J. Zhao, M. Karimzadeh, L. S. Snyder, C. Surakitbanharn, Z. C. Qian, and D. S. Ebert. MetricsVis: A visual analytics system for evaluating employee performance in public safety agencies. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1193–1203, Jan 2020. doi: 10.1109/TVCG.2019.2934603