Adapting Level of Detail in User Interfaces for Cybersecurity Operations

Catherine Inibhunu, Scott Langevin, Scott Ralph, Nathan Kronefeld
Uncharted Software Inc.
Toronto, Ontario, Canada

Harold Soh, Greg A. Jamieson, Scott Sanner, Sean W. Kortschot, Chelsea Carrasco, Madeleine White
University of Toronto
Toronto, Ontario, Canada

Abstract— As cybersecurity threats increasingly appear in news headlines, the security industry continues to build state of the art firewall and intrusion detection systems for monitoring activities in complex cyber networks. These systems generate millions of log files and continuous alerts. In order to make sense of cyber data, cyber security and system administrators review and analyze millions of logs using highly summarized views and manual cycles of click-intensive details-on-demand. This is laborious, induces cognitive overload, and is prone to errors resulting in important information and impacts not being seen when most needed. Our research focus is on developing “FocalPoint” a system that provides Adaptive Level of Detail (LOD) in user interfaces for cybersecurity operations. FocalPoint is a recommender system tailored for complex network information structures that reasons about contextual information associated with the network, user tasks, and cognitive load. This facilitates tuning cyber visualization displays thereby improving user performance in perception, comprehension and projection of current Cybersecurity Situational Awareness (Cyber SA). For cyber analysts, having the right information, in context, when most needed without cognitive overload could lead to effective decision making in cyber operations. We provide a use case scenario for FocalPoint with an in-progress prototype and highlight various challenges and potential considerations for building an effective adaptive system.

Keywords—Adaptive User Interfaces; Cybersecurity; Network Situation Awareness; Human Computer Interaction; Context-aware Reasoning; Adaptive Visualization

I. Introduction

Cybersecurity operations involve the collection of large amounts of information from a wide variety of machines interconnected in cyber networks. The ability to analyze this information, understand behaviours and present it in a manner that can be easily consumed by security analysts and system administrators is a complex problem [1]. Today's state of the art network analysis and monitoring tools provide interfaces that: 1) require laborious manual interactions; 2) provide limited ability to effectively navigate through complex networks; and 3) often fail to account for operational goals or business processes and their inherent uncertainty [2]. Efficiency and effectiveness is only achieved with highly trained experts in cybersecurity and network administration. This approach does not scale well with operational complexity, or support the rapid response required for the dynamic nature of cyber space.

To provide Cyber SA, some researchers have focused on analyzing large volumes of network flow data [3] while others look to map threats and vulnerabilities using attack graphs and networks maps [4]. Visualization systems have been suggested to better understand the results generated from analyzing network log files [1] [5], however, these systems are not able to handle complex volumes of cyber data, perform minimal or no cross correlation, exhibit poor performance when trying to reason about context changes and lack the ability to manage multiple informational displays, which leads to cognitive overload for users [1].

To increase the effectiveness of visual analytic tools, there is a need to build systems that attend to the needs and abilities of the user [6]. Such systems have been utilized for creating adaptive models that generate user tailored recommendations in web interfaces, human learning tools and desktop assistance [7].

In a dynamic cyber space where uncertainty is hard to quantify, deciding what information is most relevant to the user is a central challenge. Cyber analysts are expected to change LOD frequently from monitoring, inspection and planning. Systems are needed that assist analysts to maintain cyber SA in large complex networks by effectively managing information overload. By adaptively tailoring the LOD presented as analysts navigate back and forth from high-level context to low-level detail, human performance can be improved by increasing the scale and complexity of networks that can be monitored, and enhance decision-making by focussing on critical information [8].

This paper presents our research on adaptive LOD in visual analytic tools for cybersecurity operations. We seek to improve human performance in cyber operations by mitigating information overload and facilitating Cyber SA. FocalPoint is a real-time adaptive system for determining appropriate LOD views tailored for hierarchical network information structures. FocalPoint reasons about contextual information associated with the network, user tasks and cognitive load to tune the presentation of network displays. Our hypothesis is that this

978-1-5090-2002-7/16/$31.00 ©2016 IEEE

adaptive visualization will improve user perception, comprehension and projection of threats and vulnerabilities.

The rest of the paper is organized as follows: section II provides a summary of related work, section III describes a use case scenario with a prototype for FocalPoint, and finally conclude the paper in section IV.

II. Related Work

Our work integrates research from various fields: adaptive systems, human computer interfaces, user modeling, adaptive visualization and their application in cyber SA.

Cyber Situation Awareness

In cyber security operations, the need to maintain SA has been studied extensively. In particular, visual analytic approaches support analytical reasoning facilitated by interactive visual interfaces and integration with computational analytics. For example, detection of anomalies and inferring interesting patterns from network flow data [3]. Another approach is the use of attack graphs that formulate potential paths an attacker might exploit to compromise systems [4].

Adaptive Systems

Visualization is suggested as a way to detect and understand network vulnerabilities and threats. However, current systems are not able to scale to large networks, resulting in cognitive overload and ineffective cyber SA [4]. To overcome this, adaptive systems are suggested to utilize users' knowledge, background and knowledge of user tasks [9]. Techniques include Bayesian networks to infer user focus [2], and content adaptation, navigation and presentation [10]. Other works introduce taxonomies to identify levels of adaptation and their triggers [11], weighted simulations to evaluate tasks and goals for adaptation [12], models for minimal interaction between human and machine [13], and reasoning on possible states for adaptation using ontology based context trees [14].

Adaptive Visualization

Adaptive Visualization (AV) research aims to reduce information overload [8]. Techniques include fuzzy logic to infer context from incoming data, human concepts and real world situations to support gradual tuning of visual displays [15]; detection of user anomalies [8]; and ontologies to depict visualization paradigms [16].

User Modelling and Human Computer Interfaces

In order to adequately facilitate the breadth of tasks involved with active network monitoring, decision support systems (DSS) need to adapt to the psychological demands of the current task [17] [18]. For example, learning tasks require uninterrupted, focused attention [17] [19], whereas pattern recognition tasks require divided, distributed attention [20]. Other research has examined inferring user attentional state via a combination of biometrics and user input [21], adapting visualizations to individual user models [35], facilitating preparatory processes for attentional switching [36], and modelling visual search through ACT-R models [37].

III. FocalPoint

the user, context, and environment to accurately infer analytic intentions and needs of the user [8].

To address these challenges, we are investigating the integration of human-computer interfaces, user modeling and context-aware adaptive visualization for application in cyber SA. FocalPoint reasons about contextual information associated with the network, user task and cognitive load to tune the presentation of interactive network visualization displays. FocalPoint contains five major components [26];

Activity Monitor: This represents the initial phase of the system and involves acquiring information about the network and the user. Examples of network activity include data flow quantity, data flow type, login attempts, and anomalous events, which are significant deviations from what is expected given the current state of the network. User data describes user interaction with the system, such as window control (e.g., zoom and pan), node and edge interaction, and task specific activities (e.g., flagging anomalous behaviour). This data informs the subsequent components of FocalPoint.

Context Reasoning: To infer the state of the network and the user, this component integrates machine learning methods for context reasoning [22].

Adaptive Decision: To guide the decisions for adaptation, this component utilizes decision theoretic models drawn from graphical and probabilistic models [23] [10] [24].

LOD Adaptation Modules: Using situational assessment and triggers from context reasoning and adaptive decision, this component governs how to tailor the adaptation - what level of detail to present and the appropriate mode of adaptation (recommend or automate).

User and Task Reasoning: This component utilizes context models to infer the task or subtask the user is currently engaged in given a sequence of activity or events received from user activity sensors. Techniques used draws from work in human-computer interfaces and integrate human factors principles [25] [9] [7].

Our focus on adaptive display of network details is on three user-system interactions: 1) User requests for information, where the system retrieves the requested information and tailors the associated level of detail, e.g. query for sub-graph or navigating network map; 2) The system receives new information from network sensors or analytic services, and the current display is updated, e.g. updates to network properties such as latency, utilization, or connectivity; 3) The system reasons about the user, the context and makes recommendation on potential focus area for the user.

Below we illustrate the approach using a use case scenario that incorporates the five components of FocalPoint.

Use Case Scenario: Network Situation Awareness

Huge volumes of log files are continuously generated from firewalls and intrusion detection systems (IDS) monitoring complex cyber networks. Representative analyst tasks aim to answer these key analytical questions: Using these firewall logs, can we infer if an activity is anomalous? Using the IDS logs, can we detect if there is a threat? Can we characterize the type and extent of any threat? FocalPoint aims to assist a cyber analyst answer these questions by reasoning about the underlying context and providing appropriate levels of details without inducing cognitive overload.

network can be facilitated by using tools that anticipate the analyst's needs, assess the value of information, and reason about context. Innovative methods are needed to understand

which provided cyber logs for a fictitious organization, Bank World (BW). The data set entails millions of records from network and firewall logs and an intrusion detection system monitoring the organization over a period of 2 days [27]. The data contains approximately 4000 machines of different types (i.e. mail server, workstations, etc.) with varying importance to BW.

To facilitate **Context Reasoning**; we use two stages: 1) data pre-processing; and 2) generation of deterministic attack paths.

Data Pre-processing: To better understand the behaviors in the BW network we first analyze the firewall logs and IDS logs to: a) infer if the traffic flow is normal or not; b) if there is an attack in the network, identify the source and potential targets; and c) identify the impact of an attack. Class probabilities were generated using a Naive Bayes model based on behavior corresponding to alerts generated by the IDS system. Using this data, a network topology is generated (nodes represent machines and links represent the activity flow between machines). Fig 1 shows a subset of the network topology that contains 143 machines with unique attributes based on machine type, priority, and number of connections flowing to and from the machines. This topology allows one to quickly view the traffic flowing in and out of the BW network, which machines are receiving high traffic and identifying those that are internal or external to BW using predefined clusters.

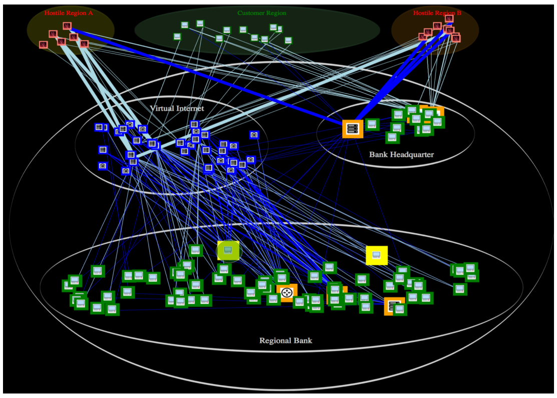

Fig 1. Sample Network Topology using Clusters: Machines in BW are grouped into 3 internal clusters (the virtual internet, the regional bank and the headquarters). There are 3 clusters of machines outside the BW cluster, which are simulated to depict 2 hostile regions and 1 region for the bank customer base. The display adapts to show traffic flow between the different regions with more emphasis using link size and node color on traffic coming from hostile regions.

Deterministic Attack paths: These are generated to infer the state of the network. A visual representation describes the potential paths an attacker might take to compromise connected machines in the network. Using the transitive inference rule [28], if a machine A has a direct connection to a machine B and a machine B has a direct connection to machine C, then we can logically infer that there is a direct link between machine A to C. Relating this to our network space, if an attacker has gained access to machine A then they will also be able to gain access to machine C through B.
We add risk assessments to nodes to identify which exhibit suspicious behaviours and those that are predicted to be next targets from a potential attacker.

information importance, we are building context aware reasoning services that adapt views of key elements in cybersecurity, including node-link visualizations of network maps, and attack graphs. We are looking to integrate adaptive user interfaces similar to work in [8] with Bayesian reasoning [29] to provide real-time information about impacts of attacks, potential severity, and mitigation strategies, without causing information overload. Specifically, we are exploring (approximate) inference using a dynamic Bayesian network (BDN) that captures the user's underlying state of knowledge conditioned upon interaction histories. The importance of informational elements can then be quantified using principled measures such as Value of Information (VOI) or convex surrogates.

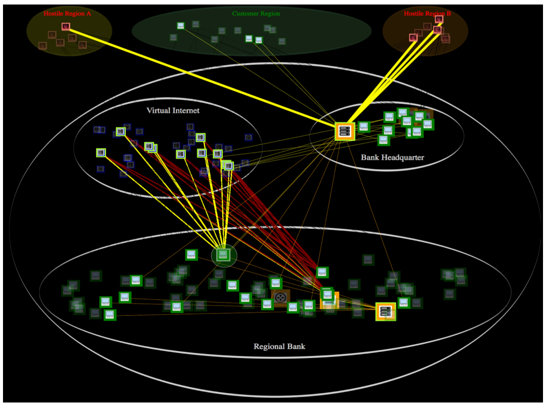

Fig 2: Simulated Attack Pattern: Deterministic attack paths show the various machines that can be compromised in a subnet. Adaptation brings focus to directly connected machines using color, size and salience.

Adaptive Display of Network Details with Adaptive Level of Detail Reasoning Service: We recognize that building a reasoning service is crucial to adapting the level of detail in displays. As shown in Fig 2, the deterministic attack paths form a graph traversal problem with connected nodes and links that are hard to understand even on small networks [4]. Preliminary displays focus on selected subnets in the network; the challenge is integrating multiple heterogeneous levels of detail. Work is currently in progress to utilize similar models as suggested in [11] to effectively identify triggers for adaptation and reason on what level of detail to display, which can vary at different locations in the display. In addition, we are exploring principles for effective adaptive visualizaton [8].

Automation and Recommendation methods for Usable Adaptations with User Models and Human Computer Interfaces: To reason about an operator's attention, behaviour, current task and cognitive state, FocalPoint will be instrumented with user activity monitoring [38]. User state can be inferred by modelling the relationship between the current system state user interaction with the system [23]. In contrast to existing work based on Hidden Markov Models [30] and Bayesian Networks [31], [32], we are investigating deep learning methods, specifically Long Short-Term Memory (LSTM) recurrent neural networks [33], which have achieved success in many domains that involve temporal data.

In preliminary work to perform both next-action prediction and task recognition in FocalPoint, we developed

Context Reasoning: In order to enable the system to make appropriate decisions on when to tailor displays based on

Tensorflow [34]. In preliminary experiments using a dataset collected from human users on a desktop task [25], the deep generative model identified tasks with 88-92% accuracy (cross-validation scores), compared to scores of 76% and 88% for a Hidden Markov and Conditional Random Field models [25]. We aim to refine these models and integrate the user task and reasoning component in FocalPoint to adapt the display in support of user tasks.

IV. CONCLUSION

The ability to understand anomalies using data from complex cyber networks can be greatly enhanced using effective visualization interfaces [3] [1]. In this paper we presented FocalPoint, a system that integrates research from various fields: adaptive systems, human computer interfaces, user modeling, adaptive visualization, and their application in Cyber SA. We argue that effective adaptation with LOD viewing can greatly increase human performance through the mitigation of information overload and therefore facilitate effective decision making in cybersecurity. We recognize the challenges for seamless integration of user, context and the displaying interface and are investigating effective methods for integration to FocalPoint.

In future work we will perform component level tests and user experimentation to measure the effectiveness of FocalPoint using analytical cyber SA scenarios and compare against a baseline.

ACKNOWLEDGEMENT

This research is a collaboration between Uncharted Software Inc and University of Toronto with support from Office of Navy Research (ONR) Command Decision Making (DM) Program under Contract N0. N00014-15-C-5033 to Uncharted Software Inc. and a Voucher for Innovation and Productivity matching grant from the Ontario Centres of Excellent.

REFERENCES

[1] G. A. Fink, C. L. North, A. Endert and S. Rose, "Visualizing Cyber Security: Usable Workspaces," in *6th International Workshop on Visualization for Cyber Security*, 2009.

[2] E. Horvitz, "Uncertainty, Action, and Interaction: In Pursuit of Mixed-Initiative Computing," *Intelligent Systems,* pp. 17-20, September 1999.

[3] D. M. Best and B. A. Cox, "Clique: Situational Awareness through Behaviour," *Smart Systems,* pp. 66-68, July/August 2015.

[4] S. Noel and S. Jajodia, "Managing Attack Graph Complexity Through Visual Hierarchical Aggregation," in *CCS Workshop on Visualization and Data Mining for Computer Security*, Fairfax VA, 2004.

[5] F. U. and B. J., "Cyber situational awareness: A systematic Review," *computers & security ,* vol. 46, pp. 18-31, 2014.

[6] G. Carenini, C. Conati, E. Hoque and et al, "Highlighting Interventions and User Differences: Informing Adaptive Information Visualization Support," in *CHI*, Paris, 2013.

[7] A. Jameson, "Adaptive Interfaces and Agents," in *Human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications*, New York, Erlbaum, 2007.

[8] K. Nazemi, C. Stab and A. Kuijper, "A Reference Model for Adaptive Visualization Systems," in *HCII*, Orlando, 2011.

[9] G. Fischer, "User Modeling in Human - Computer Interaction," *User Modeling and User-Adapted Interaction,* vol. 11, pp. 65-86, 2001.

[10] V. Zarikas, "Modeling decisions under uncertainty in adaptive user interfaces," *Inf Soc,* vol. 6, pp. 87-101, 2007.

[11] K. M. Feigh, C. M. Dorneich and C. C. Hayes, "Toward a

[12] O. Topcu, "Adaptive decision making in agent-based simulation," *Simulation,* vol. 90, no. 7, pp. 815-832, 2014.

[13] H. Soh, S. Sanner and G. Jamieson, "A Decision-Theoretic Approach for Adaptive User Interfaces in Interactive Learning Systems," 2015.

[14] E. Vassev and M. Hinchey, "KnowLang: Knowledge Representation for Self-Adaptive Systems," *Software Technologies,* pp. 81-84, 2015.

[15] P. D. Haghighi, B. Gillick, S. Krishnaswamy, M. M. Gaber and A. Zaslavsky, "Situation-Aware Adaptive Visualization for Sensory Data Stream Mining," in *Sensor-KDD*, Berlin, 2010.

[16] K. Ryabinin and S. Chuprina, "Development of ontology-based multiplatform adaptive scientificvisualization system," *Journal of Computational Science ,* 2015.

[17] B. P. Bailey and J. A. Konstan, "On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state," *Computers in Human Behavior*, 2005.

[18] C. Roda and J. Thomas, "Attention aware systems: Theories, applications, and research agenda," *Computers in Human Behavior*, 2005.

[19] S. Grossberg, "The link between brain learning, attention, and consciousness and cognition," *Consciousness and Cognition*, 1998.

[20] M. A. Ghazanfar, M. Cook, B. Tang, I. Tait and A. Alijani, "The effect of divided attention on novices and experts in laparoscopic task performance," *Surgical Endoscopy*, 2015.

[21] E. Horvitz, C. Kadie, T. Paek and D. Hovel, "Models of attention in computing and communication," *Communications of the ACM*, 2003.

[22] A. Abbas, L. Zhang and S. U. Khan, "A suvey on context-aware recommender sytems based on computational intelligence techniques," *Computing,* vol. 97, no. 7, pp. 667-960, 2015.

[23] A. Jameson and B. G.-H. e. al., "When actions have consequences: Empirically based decision making for intelligent user interfaces," *Knowledge-Based Systems,* vol. 14, no. 1-2, pp. 75-92, 2001.

[24] S. Abraham and S. Nair, "A Predictive Framework for Cyber Security Analytics using Attack Graphs," *International Journal of Computer Networks & Communications ,* vol. 7, no. 1, 2015.

[25] A. Elbani and M. N. Omri, "Web User Interact Task Recognition Based on Conditional Random Fields," *Computer Analysis of Images and Patterns,* vol. 9256, pp. 740-751, 2015.

[26] C. Inibhunu and S. Langevin, "Adaptive Visualization of Complex Networks with FocalPoint; A context aware level of detail recommender system," in *HFES*, Washington DC, 2016.

[27] VAST., *http://www.vacommunity.org/VAST Challenge,* 2012.

[28] J. Rauch, "Logic of Association Rules," *Applied Intelligence,* vol. 22, no. 1, 2005.

[29] F. Jensen and T. Nielsen, Bayesian networks and decision graphs, NY: Springer, 2007, p. 269–282.

[30] J. Shen, "Activity recognition in desktop environments," Ph.D. dissertation, Oregon State University, 2009.

[31] S. Koldijk, "Look what you've done! Task recognition based on PC activities," M.S. thesis, Radboud University, Nijmegen, The Netherlands, 2011.

[32] V. Magnanimo, M. Saveriano, S. Rossi and D. Lee, "A bayesian approach for task recognition and future human activity prediction," in *International Symposium on Robot and Human Interactive Communication*, 2014.

[33] S. Hochreiter and J. Schmidhuber, " Long short-term memory," *Neural computation,* vol. 9, no. 8, pp. 1735-80, 1997.

[34] A. Agarwal, P. Barham, E. Brevdo, Z. Chen and C. C. G. S. .. .. .. Z. X. Citro, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," 2015.

[35] D. Toker, C. Conati, G. Carenini and M. Haraty, "User Modeling, Adaptation, and Personalization," Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012.

[36] A. Kiesel, M. Steinhauser, M. Wendt, M. Falkenstein, K. Jost, A. Phillip and I. Koch, "Control and interference in task switching—a review," *Psychological Bulletin*, 2010.

[37] M. Fleetwood and M. Byrne, "Modeling the visual search of displays: A resived ACT-R model of icon search based on eye-tracking data," *Human-Computer Interaction*, 2006.

[38] D. Schroh, N. Bozowsky, M. Savigny and W. Wright, "nCompass