CyberCopter: a 3D helical visualisation for periodic signals of cyber attacks

Nicolas Delcombel, Alexandre Kabil, Thierry Duval, Marc-Oliver Pahl

CyberCopter: a 3D helical visualisation for periodic signals of cyber attacks

Delcombel Nicolas
IMT Atlantique, Brest, France
Lab-STICC, UMR CNRS 6285

Kabil Alexandre
IMT Atlantique, Brest, France
Lab-STICC, UMR CNRS 6285

Duval Thierry
IMT Atlantique, Brest, France
Lab-STICC, UMR CNRS 6285

Marc-Oliver PAHL
IMT Atlantique, Rennes, France
IRISA, UMR CNRS 6074

Abstract

This paper [1] aims to asses the usefulness of 3D interactive interfaces to display periodic signals in a network. Past research has shown that 2D data visualization simplifies alerts classification. Including those drawn by periodicity based Intrusion Detection Systems. However, 2D visualisations have limitations such as screen space availability. This is why we created CyberCopter, a prototype that uses a 3D helical representation to highlight periodic patterns in a dataset. We tested CyberCopter usability and efficiency in a fraud detection scenario. It scored 77 at the SUS questionnaire which demonstrates an acceptable usability.

1 Introduction

Networks are growing in size and so is the number of attacks they receive. These attacks leave periodic signals that can be detected. Interactive visualizations help operators understand these signals. However, these 2D visualizations are limited [11], for example in display space or in the interactions they offer.

A solution to solve the problem of data representations is Immersive Analytics (IA). IA explores how new interactions and display technologies can support analytical reasoning and decision making. It can be used to provide new ways to explore complex data [4].

Indeed, 3D tools meant to help operators in their tasks can also speed up their training [10]. These Information-Rich Virtual Environments (IRVE) help humans organize their data, because spatialization improves memory and recall [16], increasing training efficiency.

To try to respond to the following question: "How can Immersive Analytics be useful to cybersecurity?", we present a case study of how IA is applied to investigate periodic alerts in a network. In section 2, we present the related work on interactive cybersecurity visualizations, periodic alerts visualization in 2D, and periodic signals representation in 3D. In section 3, we describe our prototype : the CyberCopter. In section 4, we present the experiment we conducted and the results obtained. In Section 6, we present our projects for future work.

2 Related Work

In this section, we present the existing solutions for the exploration of temporal signals of the network following an alert raised by an Intrusion Detection System (IDS). Generally, agents use an IDS that raises alerts but does not allow them to explore the network. So we focus on research on interactive visualizations that aims to support agent's investigation of an alert. We specifically focus on 2D representation that aims at displaying periodic patterns created by attacks. 3D representations for periodic signals in cybersecurity and other domains encourage us to think that Immersive Analytics can overcome their limitations.

Most malwares leave periodic traces in the network traffic [1]. However, healthy systems also produce periodic signals. Therefore, using signal processing to detect the presence of malware generates many false positive alerts [9]. While some researchers try to find a robust way to detect hidden periodic signals in network traffic, visualization interfaces focused on highlighting periodic patterns that have emerged [8] [9] [5].

[1] This work was done at the chair Cyber CNI with support of the FEDER development fund of the Brittany region, France.

USENIX Symposium on Usable Privacy and Security (SOUPS) 2021.
August 8–10, 2021, Vancouver, B.C., Canada.

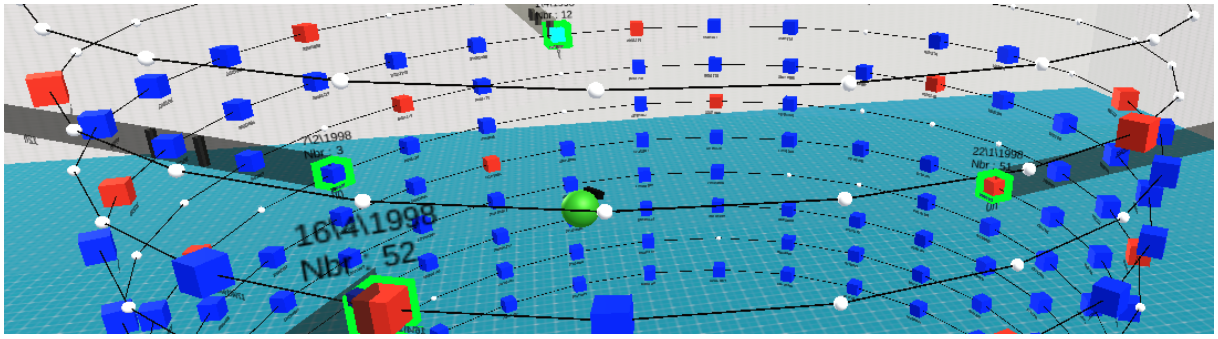

Figure 1: External view of CyberCopter. The green ball represents the position where a user should be to benefit from an internal view of the helix.

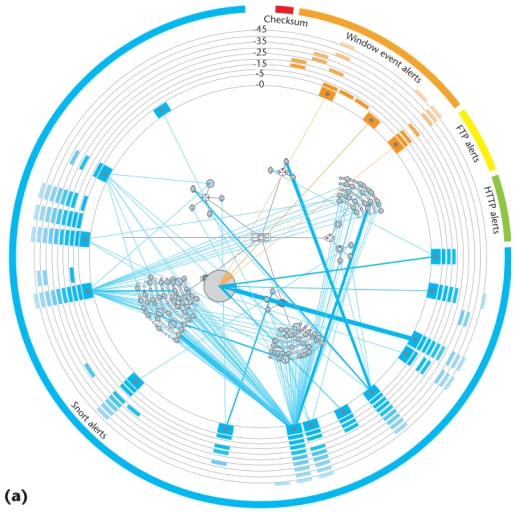

Figure 2: Alerts are displayed on concentric circles, depending on the time spent since they have been received. Their angular position and color depends on the type of alert (windows, snort ...).

2.1 2D Representation of Periodic Signals for Cybersecurity

2D interfaces usually follow guidelines that can be used to improve 3D cybersecurity representation. 2D interfaces that highlight periodic signals are interactive and can help represent the context of an attack raised by an IDS [7]. They facilitate the investigation process by displaying information from multiple sources [11].

Interactive representation should follow Shneiderman's mantra [15]: Overview first, next zoom and filter, and then present details on demand. Shneiderman adds three more requirements: view relationships among items, have a history of actions, and extract data. 3D representations should follow these requirements as well to be as usable as possible.

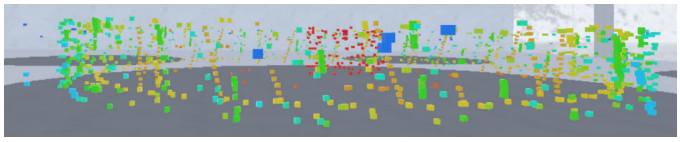

Figure 3: Helovis : Cubes color and position on the radius represent the frequency of the signal. Their position depends on the time of reception. Participants can modulate the period of the helix to make patterns appear.

While Gove et Deason [8] use a histogram to represent the temporal aspect of data, Huynh et al. [9] use spiral to highlight periodic patterns inside the dataset. A spiral period is the number of steps between two adjacent points on the radius. By modifying the spiral's period, one can easily spot periodic patterns in a dataset [17]. Foresti et al. [5] display alerts on concentric circles, the further from the center, the oldest an alert is. This representation allows to spot a repetition of alarms in time like spirals, but the user can't change the periodicity displayed. This representation leaves space at the center of circles to display where the attacks represented happened on the network (figure 2°.

These representations have limitations: lines are covering the rest of the representation, and spirals representations take more space than histograms or concentric circles. It leaves less space to display other dimensions of the dataset, which could have been used to give an idea of the context.

2.2 3D Representation of Periodic Signals

3D representations offer more space to display additional information and link them without occlusion [13]. A helix has the advantages of being able to change periods like a spiral, and to leave enough space to other information like concentric circles. Scott et al. [14] use a helix to represent attacks on a system during a day. Additional parameters such as attack severity and frequency, IP of the receiver, can be

mapped to the node placement on the radius of the helix. To our knowledge, one of the most recent research on helix visualization in virtual reality is by Cantu and al. [3] in the field of radar signal detection (figure 3). Based on Gestalt laws and depth perception, HeloVis makes possible to detect periodicity of radar signals easier than with usual 2D systems. Moreover, this representation is robust to missing data and helps to ignore the noise from the signals. In the case of Helovis, Virtual Reality helps to interact with data easily and intuitively.

2.3 Synthesis

Although they have some limitations about display space, what is interesting with 2D visualizations is that they follow Shneiderman's mantra and link additional information to the temporal representation.

3D visualization seems able to overcome these 2D limitations and the helix seems to be the best representation in 3D to highlight temporal patterns in data.

This is why we want to take the best of these two worlds to propose a new 3D approach to ...

3 Prototype

Based on the related work, we created a prototype (figure 1) that uses a helix to display periodic alerts in virtual reality. The 3D organisation allows the representation to respect the Shneiderman's mantra. The participant can always see the helix, providing him the context. S/he can filter specific data points and display more information on demand.

To test its usability in simple scenarios, we used the 100k Movies dataset. This dataset is used by Webga et al. [18] to simulate a fraud attack on movie ratings. Because fraud detection and breach detection have many similarities, this dataset is usable for our use-case. It consists of grades given to movies by different users. The dataset includes more than 100,000 votes by 943 users on 1682 time-stamped movies over a period of 7 months (October 97 to April 98). Each user rated at least 20 movies and the ratings are integers between 1 and 5 inclusive. Also, for each vote, information is given about the user: age, gender, professional situation, zip code. For their experimentation, Webga et al. added fictional users who artificially changed the rating of a movie. The aim of the system they presented was to detect these users. We used a similar scenario to test our prototype.

In the prototype [2], each cube represents a day. The color of the cube (dark blue, light blue, green, yellow, or red) signals the number of votes this day. The helix period can be changed with a slider at the bottom right of the User Interface (UI). The UI contains information about the selected movie: id and number of users who voted for this movie. To explore the scene, the participant moves with keyboard arrows and looks around with the mouse.

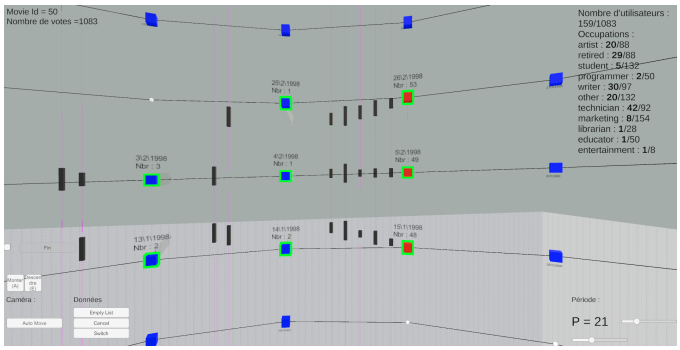

Figure 4: Complementary information appears when the user clicks on a cube. The position of the parallelepiped represents the mark given, the closer to the cube meaning 1/5 and the further 5/5. Its size represents the number of votes for this mark.

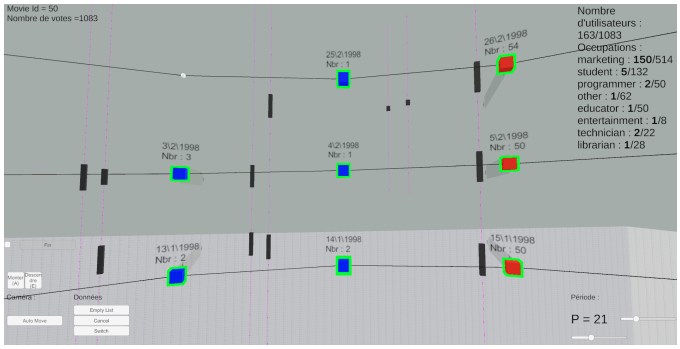

Figure 5: In this case, an unusual can pattern can be discerned between cubes with high number of votes compared to Figure 4. Almost all the marks given are 1/5.

When clicking on a cube, additional information appears behind the helix (figure 4). It can represent the distribution of votes per hour or marks given during the day selected. The participant can choose which information to display by using the "switch" button on the interface. This representation helps the participant compare the data of different days together (figure 5).

Selecting data updates the list at the right of the screen. This list shows the occupation of users who voted on the selected days. The participant can use the cancel button to reset the list and have access to the information of all the users who voted for the movie. S/he can use the empty button to empty the list if S/he wants to have information on another selection of users. S/he can change the angle between the radial information and the helix using a slider on the UI.

This representation uses the space beyond the helix to add

[2] Available at: https://r0rchard.github.io/CyberCopter/

information. Additionally, radial representations are linked to their respective days on the helix using their positioning in 3D space. This method does not use lines to link together multiple representations. This helps the visualisation to be less obstructed and easier to understand at a glance.

4 Experiment and Results

Our primary focus is to test the prototype usability. Participants recruited are not cybersecurity experts. We chose a simple scenario for them to understand in a short amount of time. This scenario is validated by our industrial partners as being representative of cybersecurity scenarios.

Because of the current health crisis, the experiment took place in telepresence on the participant's computer web browser. Participants used their mouse and keyboard to interact with the 3D environment. Despite the variability of systems, there was no performance issue. After signing an online consent form, the participant learns about the application in a tutorial. During the experiment, experimenters have access to the participant's actions and can help him/her if he/she encounters a problem that does not fall within the scope of the experiment (e.g. a bug).

Participants followed this scenario : an alert is raised on a specific movie, in order to know if it is a false alarm or not, they investigate the movie's data using cybercopter. Malicious users have common characteristics and behaviors: they all have the same job, and they all rate at the same time, with the same mark which suggests fraud. The scenario ends when the participant judges that a fraud is present or not.

After the experiment, s/he answers the different questionnaires via an online form. Completion time is measured as well as the usability of the prototype using the System Usability Scale questionnaire [2]. Additionally, The following data are collected: each click on a cube, each click on a button, and which button was clicked. This allows us to know how the participant searches for information, how many times he/she goes back and which features are used or ignored. We hope to better understand their reasoning in 3D space.

For now, 8 participants aged between 23 and 30, 5 women and 3 men, have taken part in the experiment. Seven of them raised the alarm with correct insight in less than 10min, the last one took longer (40 min) and was the only one without gaming experience. While the number of participants is still low, the mean SUS score of 77 is an encouraging indicator of the usability of the cybercopter.

Moreover, we received positive feedback from our industrial partners who value the possibility of adding information next to the helix.

5 Conclusion

In this paper, we propose a helical representation of data aimed at cybersecurity. The helix helps detect periodic signals in a dataset. The space left by the helix is used to display additional information. Early experiments indicate that the prototype has an acceptable usability. Our future work includes the comparisons of the prototype with 2D representation.

6 Future Work

We want to apply our prototype on a cybersecurity use-case. The SWaT dataset [6] contains data of sensors of a water treatment plant. These captors usually display periodic activity. In this dataset, different attacks disrupt the normal operation of the plant.

Lohfink and all [12] represent SWaT data with multiple spirals. In their experiment, they ask the participant to classify alerts generated by an IDS. S/he can use the representation of multiple sensors to validate if an alert raised by the IDS is indeed an attack. A 3D representation with multiple helix can improve participants situational awareness. It will allow them to compare multiple sensors together or current data with past data of an attack. Helix positioned to represent the 3D location of the captor allows participants and to understand which process is under attack and postulate the goal of the attacker. Additionally, immersion offered by VR headsets helps users engagement, adoption and satisfaction, perhaps at the expense of time of completion.

Moreover, we want to asses Cybercopter capacity to help process retention. Indeed, 3D environments help recognition and recall tasks of processes [16]. It could be used for operators training. To assess the positive effect of our prototype, we will conduct a long-term experiment.

References

[1] Ngoc Anh Huynh, Wee Keong Ng, Alex Ulmer, and Jorn Kohlhammer. Uncovering periodic network signals of cyber attacks. *2016 IEEE Symposium on Visualization for Cyber Security, VizSec 2016*, 2016.

[2] John Brooke. SUS: A 'Quick and Dirty' Usability Scale. *Usability Evaluation In Industry*, (November 1995):207–212, 2020.

[3] Alma Cantu, Thierry Duval, Olivier Grisvard, and Gilles Coppin. HeloVis: A Helical Visualization for SIGINT Analysis Using 3D Immersion. *IEEE Pacific Visualization Symposium*, 2018-April:175–179, 2018.

[4] Tim Dwyer, Benjamin Bach, Raimund Dachselt, Sheelagh Carpendale, Christopher Collins, and Bongshin Lee. Immersive analytics: Exploring future interaction and

visualization technologies for data analytics. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces: Nature Meets Interactive Surfaces, ISS 2016*, pages 529–533, 2018.

[5] Stefano Foresti, James Agutter, Yarden Livnat, Shaun Moon, and Robert Erbacher. Visual Correlation of Network Alerts. *IEEE*, (8):1275–1279, 2006.

[6] Jonathan Goh, Sridhar Adepu, Khurum Nazir Junejo, and Aditya Mathur. A dataset to support research in the design of secure water treatment systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10242 LNCS(October):88–99, 2017.

[7] John R. Goodall, Wayne G. Lutters, Penny Rheingans, and Anita Komlodi. Preserving the big picture: Visual network traffic analysis with TNV. *IEEE Workshop on Visualization for Computer Security 2005, VizSEC 05, Proceedings*, pages 47–54, 2005.

[8] Robert Gove and Lauren Deason. Visualizing Automatically Detected Periodic Network Activity. *2018 IEEE Symposium on Visualization for Cyber Security, VizSec 2018*, pages 1–8, 2019.

[9] Ngoc Anh Huynh, Wee Keong Ng, and Hoang Giang Do. On periodic behavior of malware: Experiments, opportunities and challenges. *2016 11th International Conference on Malicious and Unwanted Software, MALWARE 2016*, pages 85–92, 2017.

[10] Alexandre Kabil, Thierry Duval, Nora Cuppens, Gérard Le Comte, Yoran Halgand, and Christophe Ponchel. 3D CyberCOP: A collaborative platform for cybersecurity data analysis and training. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11151 LNCS:176–183, 2018.

[11] Philip A. Legg. Visualizing the insider threat: Challenges and tools for identifying malicious user activity. *2015 IEEE Symposium on Visualization for Cyber Security, VizSec 2015*, (October 2015), 2015.

[12] Anna Pia Lohfink, Simon D.Duque Anton, Hans Dieter Schotten, Heike Leitte, and Christoph Garth. Security in Process: Visually Supported Triage Analysis in Industrial Process Data. *IEEE Transactions on Visualization and Computer Graphics*, 26(4):1638–1649, 2020.

[13] Arnaud Prouzeau, Antoine Lhuillier, Barrett Ens, Daniel Weiskopf, and Tim Dwyer. Visual Link Routing in Immersive Visualisation Arnaud. *28th Modern Artificial Intelligence and Cognitive Science Conference, MAICS 2017*, pages 189–190, 2017.

[14] Craig Scott, Kofi Nyarko, Tanya Capers, and Jumoke Ladeji-Osias. Network intrusion visualization with niva, an intrusion detection visual and haptic analyzer. *Information Visualization*, 2(2):82–94, 2003.

[15] Ben Shneiderman. Eyes have it: a task by data type taxonomy for information visualizations. *IEEE Symposium on Visual Languages, Proceedings*, pages 336–343, 1996.

[16] Richard Skarbez, Nicholas F. Polys, J. Todd Ogle, Chris North, and Doug A. Bowman. Immersive Analytics: Theory and Research Agenda. *Frontiers in Robotics and AI*, 6(September), 2019.

[17] Christian Tominski and Heidrun Schumann. Enhanced Interactive Spiral Display. *The Annual SIGRAD Conference Special Theme: Interaction*, (May):53–56, 2008.

[18] Kodzo Webga and Aidong Lu. Discovery of rating fraud with real-time streaming visual analytics. *2015 IEEE Symposium on Visualization for Cyber Security, VizSec 2015*, pages 1–8, 2015.