CyberPetri at CDX 2016: Real-time Network Situation Awareness

Dustin Arendt, Dan Best, and Russ Burtner
Pacific Northwest National Laboratory

Celeste Lyn Paul
Department of Defense

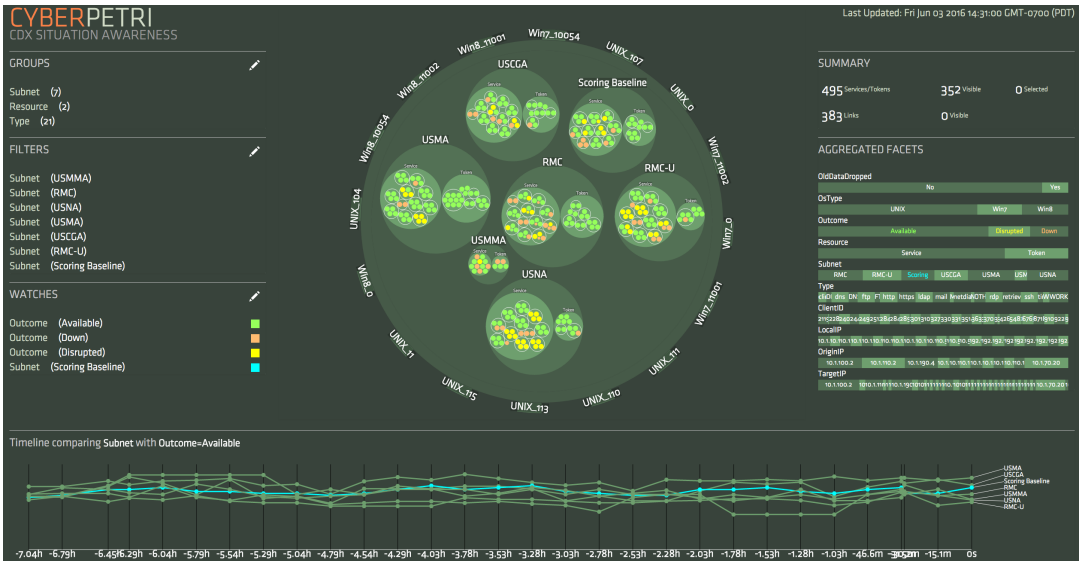


Figure 1: CyberPetri showed the status of services and tokens across the participating academies live during CDX 2016.

Abstract

CyberPetri is a novel visualization technique that provides a flexible map of the network based on available characteristics, such as IP address, operating system, or service. Previous work introduced CyberPetri as a visualization feature in Ocelot, a network defense tool that helped security analysts understand and respond to an active defense scenario. In this paper we present a case study in which we use CyberPetri to support real-time situation awareness during the 2016 Cyber Defense Exercise.

Keywords: Cybersecurity; situation awareness; visualization

Index Terms: C.2.3 [Computer-Communication Networks]: Network Operations —network management, network monitoring

1 Introduction

The Cyber Defense Exercise (CDX) [3] is a Red vs. Blue cybersecurity competition that tests how well participants can secure and defend their networks. Multiple Blue teams configure and secure a virtual network that hosts services such as websites, email, and remote access. A Red team then spends four days attempting to gain access to data and services on the Blue teams' networks without being detected. The Blue teams must detect and defend against Red team attacks while maintaining normal network services. Points are awarded to the Blue team for maintaining service availability and for protecting the confidentiality, integrity, and access of data from the Red team. The Blue team with the most points at the end of the competition wins. A neutral White team keeps track of Blue team scoring and ensures that both Blue and Red teams play by the rules.

Blue team participation is open to all U.S. military service academies, which included the U.S. Military Academy (USMA), Naval Academy (USNA), Coast Guard Academy (USCG), and Merchant Marine Academy (USMMA), as well as graduate and undergraduate teams from the Canadian Royal Military College (RMC and RMC-U). The Red and White teams were volunteers from the service academies as well as cybersecurity professionals in the Department of Defense.

In this case study, we describe how we redesigned the CyberPetri visualization technique for real-time network situation awareness. First we describe how we adapted CyberPetri to support CDX 2016, including data preprocessing methods, application architecture, and user interface design. Then in the case study, we describe useful visualization configurations, situation awareness successes, and general usability feedback from users. Finally, we discuss lessons learned from CDX and plans for future work.

2 CyberPetri

We redesigned the Ocelot research prototype [1] to support situation awareness at CDX. In previous exercises, the White team relied on simple descriptive charts (e.g., bar charts, line charts, pie charts) of summative scoring results retrieved from the scoring database. However, this view of the network was not detailed enough to convey status about individual machines and services. We designed CyberPetri to provide a more nuanced view by visually encoding services or tokens running on an individual machine as the leaf nodes of the circle packing visualization. Color was used to communicate the availability the services and integrity of the tokens.

2.1 Data Preprocessing

Scoring agents that monitored network services and data tokens were installed on all Blue team computers to support scoring. Events returned by these agents were aggregated into XML log files made available in batches every 15 minutes. The logs for network services and data tokens had slightly different schema and were logged to separate files. For both types we extracted the time, origin and local IP addresses, service type, operating system, and outcome (success/failure). We enriched the data by adding a field for subnet

978-1-5090-1605-1/16/$31.00

(found using a lookup table), a unique identifier for the event (found by concatenating the origin IP, local IP, and service/token type), and a flag for the resource (either service or token). We refer to a unique (Local IP, Origin IP, type) triplet within a log file and its associated metadata as an "event." The scoring agents reported service availability every 5-30 seconds for most services and data token validity every 50-70 seconds. As a result, the raw log files contained service and token event outcomes that fluctuated between success and failure many times within a batch log, thus the last known state in the log would not provide effective situation awareness. To solve this problem we wrote simple analytics to summarize these changes.

If a service had "Outcome=SUCCESS" in 90% or more of the events in the batch log, we replaced the outcome with "Available." If the service was reported as "Outcome=SUCCESS" in 10% or less, we replaced the outcome with "Down." For the remaining cases that occur between a 10% to 90% success rate, we replaced the outcome with "Disrupted." The case for data tokens was different. Knowing if a token was compromised within a period of time was more important than knowing how long it was protected. So, if a token was reported being not available in a batch at least once, we replaced the token outcome with "Down." By enriching and making the schemas for tokens and services consistent we were able to fuse the datasets and show then together in the visualization.

Sometimes the scoring agents would cache old events and include them in a current log batch. We believed including these stale results in the current visualization would be misleading. So, we discarded service events more than 16 minutes older than the latest service event and token events more than 30 minutes older than the latest token event. These thresholds were derived from the expected rate that these log files were made available.

2.2 Application Architecture

CyberPetri is a web based visualization system consisting of three components: a data processor/adapter, a web server, and the web client visualization. The data processor listens for new batches of scoring data to be written. When it becomes available the file is processed (see section 2.1). The newly processed data is posted to the web server. The web server manages an in-memory model of the processed data and hosts the necessary web resources for the client application. The client application synchronizes its model with the server's model allowing newly processed data to propagate to connected clients efficiently. When new data is available to the client, the visualization incorporates this data.

2.3 User Interface

The CyberPetri visualization (Fig. 1) is dominated by a circle packing view of the data. Concentric circles visually represent a hierarchy of results via physical containment. Containing circles organize the results in different ways depending on the fields (and their order) selected by the user. The user can further configure the appearance of the visualization through filters.

A filter is defined by a $(key, value)$ pair, and an event matches a filter if the field in the event with the corresponding key has exactly the same value. CyberPetri has two types of filters: view filters and color filters. View filters allow the user to focus the visualization on a desired subset of the data. If one or more view filters are specified, results that match any of the view filters are shown. If no view filters are specified, all results are shown. Color filters assign colors to the circles representing results that pass that filter. However, results can pass more than one filter, so color filters are assigned a priority, and a result gets the color of the highest priority filter it passes. The user sorts the color filters to set their priority. CyberPetri also contains two supporting faceted views: a timeline and horizontal stacked bar charts. The bar chart shows the counts of key-value pairs faceted by key (see Figure 2). The faceted timeline is configured with a single group that determines the facets, and a filter that is used to aggregate the data. The timeline shows (over time) the percentage of results that passed the filter with each facet for the current view.

CyberPetri inherits from Ocelot the ability to represent additional information on an external ring surrounding the circle packing view, and to connect elements in the internal and external ring using integrating lines. Initially we intended to use this feature to show where the Red team was focusing their attention. However, due to data access constraints at CDX, we did not have live access to the pcap or NetFlow data generated by the Red team. Instead we used the outer ring to represent the various error codes returned by a scoring agent during a failure. Events containing failures were linked to the corresponding error codes on the external ring. This was intended to help explain why the failure occurred, and possibly give insight about the type of attack being used by the Red team. Edge routing [2] is useful to mitigate over plotting of edges over nodes in node-link diagrams. We developed a custom routing heuristic using the shortest path of the Delaunay triangulation of the obstacles. This was intended to decrease the amount that links plot over the circle packing view.

When new token and service data becomes available, the visualization updates, and changes in position and size (e.g. of the circles) were animated to make them easier to follow. However, the colors of circles may also change, which can lead to change blindness. We mitigated this by adding a thick white stroke around circles whose color changed after an update. The stroke's opacity gradually becomes invisible five seconds after it appears. This provides the user some time to scan the visualization and spot differences they may not have directly observed to change.

3 Case Study: CyberPetri at CDX 2016

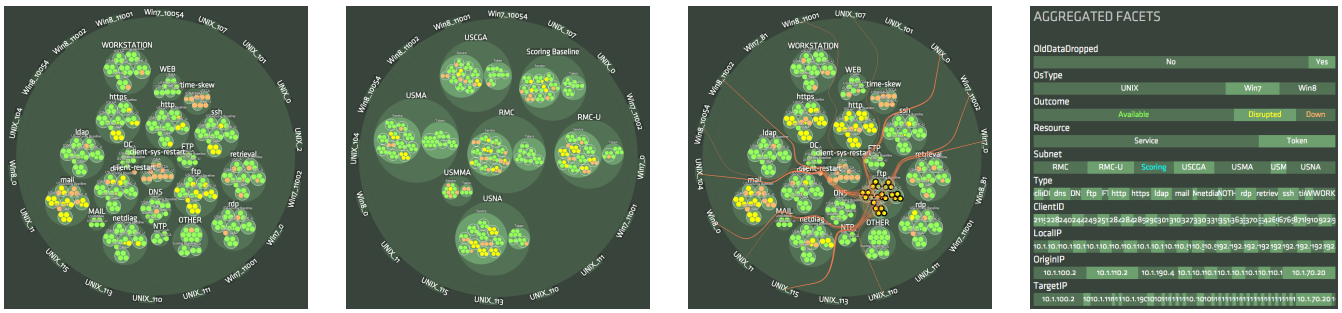
CyberPetri is configurable and agnostic to data content. However, not every configuration supported situation awareness equally well. During the exercise we relied on one configuration (see Fig. 1) that we found to be the most useful for bringing users and visitors up to speed about the happenings at CDX. We typically grouped by a permutation of Subnet, Resource and Type. We used filters to show only events related to the service academies and the scoring baseline. We colored the circles according to status so that Available events were green, Disrupted events were yellow, and Down events were red. The timeline was configured to facet the data by subnet, and show the percentage of Available events (per subnet) over time.

We found that this configuration was useful because it quickly revealed which subnets (service academies) were experiencing network disruptions or had tokens compromised. Through the drag and drop feature, we were able to quickly rearrange the ordering to allow the visualization to answer slightly different questions. For example, by moving Type to the top of the group by list, we were able to reveal which services were experiencing more disruption.

3.1 Feedback from "walk-up" users

CDX received hundreds of visitors over the course of the week. CyberPetri was available for viewing by visitors in the main lobby of the event location, in addition to a short demonstration on the official tour. Many visitors approached us with questions and comments. The majority of comments were positive, discussing how the viewer liked a particular component of the interface. However, users' questions seeking clarification and the comments mentioning potential improvements revealed ways to increase comprehension by novice or impromptu users.

A common theme emerged regarding the timeline. While the timeline view allowed for users to see the general pattern of the exercise, it was difficult for the viewers to understand more without interaction. Because the timeline used the same color for all lines, when they overlapped it was easy to lose track of the team being looked at. This was alleviated, in part, by mouse hover, which highlights the team being hovered over in white. Changing default color



(a) (b) (c) (d)

Figure 2: CyberPetri can be grouped differently to show alternate views of the data. (a,b) Results are grouped by Type and Subnet. (c) Links are routed around the large circles to mitigate over plotting. (d) Horizontal stacked bar charts show $(key, value)$ counts.

to help distinguish each line could assist with this finding, although given an arbitrary number of lines and the ability of users to color encode values of their choosing complicates the visual approach.

The outer circle of error codes also came up several times with viewers. This was not because of clarity issues, but rather because viewers did not understand what they meant. Some viewers looked up the codes (on their personal devices), while others were unsure what the values meant at all. While the codes were chosen due to lack of more meaningful data, the comments from viewers highlight the need to provide context to the information being viewed. This can potentially be accomplished through tooltips, help dialogs, or clarifying text on the display itself.

Some viewers had ideas for how to align CyberPetri with their expected use, including the following. **1) Change what a node is:** viewers requested the ability to change what a node represents. Instead of a pre-defined identifier, allow users to choose other features to be the nodes shown on the display. **2) Group toggle:** Much of the information that users were seeking was about teams. Viewers felt that allowing toggling of groups without opening the dialog would speed up their information gathering process. **3) Details collapse:** When there are many elements on the left hand side of the display, information goes off screen requiring scrolling. Instead, allowing fields to collapse would allow a user to determine which features had space on the screen.

Finally, some users provided information about the usefulness of the tool in the particular setting that was unknown previously, but could enable greater impact. One in particular was a comment that the tokens were likely of more importance than the services. This was due to the fact that the red team did not want to take down the services so that they could gather more tokens from the blue team's systems. Until then we had focused on services, but we adapted CyberPetri during the competition to more explicitly show tokens by adding the "Resource" field to events to differentiate between services and tokens. This allowing the visualization's circle packing algorithm to spatially separate services from tokens.

Overall, the interactions with the viewers who visited the display were an important component to understanding how CyberPetri could be leveraged, what some of the weak points were, and which elements were viewed as most useful.

3.2 Discussions with CDX Red & White Team

We had several impromptu interactions with the white team. On one instance, someone noticed a down service (a team having connectivity issues) on CyberPetri and then went to talk to the white team about it. The white team was not aware of the issue yet; however they were able to confirm that the team was in fact having issues. Also, a down token was noticed on the scoring baseline subnet. After discussing the token with the white team it was determined that the red team had stepped out of bounds. The white team then went to talk to red team to ensure this wasn't repeated. We had in depth discussions with two white team members who visited (at two different times) to view CyberPetri and to discuss its use, findings, and potential next steps. We also received feedback from a red team member to help verify what CyberPetri was showing. We solicited this feedback during rare periods where there was a lull in activity during the exercise.

3.2.1 White Team Member 1 (W1)

W1 was impressed by the appearance of the tool and immediately saw potential for its use. They took the time to go over CyberPetri and its features, and discuss the use of the tool in their space. W1's initial feedback was that CyberPetri could be used for white and gray team coordination and workflow. Because white and gray need to often work together to alleviate issues that teams are facing, having a tool to help coordinate that activity would speed up that process. Other insight from W1 involved the data being presented. All of the high level data presented in CyberPetri is available to the white team, however CyberPetri made that easier to find. However, the white team has access to more low-level information; W1 felt that confirmation of an event would be difficult in the tool.

3.2.2 White Team Member 2 (W2)

Interaction with W2 was shorter due to the flurry of activity happening (even late in the evening). However, they were still able to provide valuable input regarding the capability. First, confirming W1's assessment, W2 felt that having greater detail available in the tool about the information being shown would help the usability during the event. Secondly, W2 mentioned that the services should be split between incoming and outgoing to help with diagnosing issues. CyberPetri had the services shown as one node, which made it difficult (without other information) to see if disruption was caused by information going in or out of the team's infrastructure.

3.2.3 Red Team Member 1 (R1)

R1 had two comments, first that it had been a great day (Tuesday); that the red team had "gotten" nearly every school. While looking at the timeline and correlating down times to red team activities, he noted that one of the schools had shown up all day, but their connectivity was minimal. This could be due to the measure CyberPetri used to show "available" or that connectivity was more complex of an issue than overall service status reports from the scoring engine.

4 Discussion

By participating in the live exercise, we learned lessons beyond our initial findings on archival CDX data. These lessons drove new requirements and designs for future versions of CyberPetri.

4.1 Lessons Learned

The ordering of the data was a challenge that, in retrospect, was not adequately handled. Records within the log files produced by the scoring agents were time stamped, but on occasion some records' time stamps were older than the 15 minute window represented by the log file. This could occur, for example, during a network outage, which would cause events to be reported with a significant delay. We handled this case by simply throwing out these events, but a better solution would have been to update the historical model with this new data, and indicate to the user that this has occurred. The visualization should have communicated the timestamp of the data and when that data was made available to the user.

The use of the "Outcome" field for events was overloaded in a way that we didn't initially apprieciate. For tokens and most services, a "SUCCESS" or "FAILURE" value was intuitive. However, in some cases, "Outcome" was overloaded to act as a flag for certain rare events that may occur on the network, including restarting the machine. When a machine was restarted, a record was logged with "client-sys-restart:Outcome=FAILURE." However, there were no records recorded with "client-sys-restart:Outcome=SUCCESS" to indicate a time period when the machine wasn't restarted. Our missing data handler used a fill-forward rule, causing "client-sys-restart:Outcome=FAILURE" states to persist incorrectly. In the visualization this manifested as an artifact where when grouping by Type, "client-sys-restart" and "time-skew" incorrectly contained only failure states. Users frequently spotted this anomaly and asked about it. A better solution would have been to not fill forward event state, but instead to indicate missing data where appropriate.

We also found that our coloring and filtering implementations were too primitive for some use cases. For example, it was not possible to configure the filters to show only the failed tokens from a particular subnet, which is a useful sub-view of the data that was requested by some of the users.

4.2 Redesign

We have continued to evolve the CyberPetri interface based on lessons learned and for new use cases we intend to support. For these future use cases, scalability is a major concern (see Fig. 3). We are designing our interface to support interactive exploration for millions of records, which we accomplish by aggregating records in the bottom level of the hierarchy based on user-defined color filters. Abstraction of entities or records as a summary instead of individual entities creates new challenges in user comprehension. This is partially addressed in the new design through more complex filtering capabilities, where records are matched against combinations of key/value pairs allowing the user to focus on items of interest in more targeted subsets.

The new design contains an interactive timeline where brushing selects the data subset to be shown in the circle packing view. Time is broken down to 15 minute increments over the last 24 hours. This was derived from user interviews of subject matter experts who are interested in real time situation awareness. The timeline shows the count of events matching each color filter for the time increment. With color filters configured appropriately, the user can filter out normal behavior and look for anomalies that would appear as spikes or dips in the timeline. These changes support the feedback we received during the CDX exercise and new use cases on streaming data, change detection and alert management.

Figure 3 shows how our new design was used to better answer typical questions we received from participants in CDX. We configured the interface to show events grouped by Subnet and Type. Next we filtered out "Resource=Service," "Type=WORKSTATION," and "Outcome=AVAILABLE" events to show only the most critical events. The timeline shows when these events occurred within the last several hours. Brushing on the timeline focuses the circle packing view on those events only occurring within the user-defined window. From this we can easily see which teams experienced certain critical breaches during the final few hours of the exercise.

Figure 3: The new design of CyberPetri, based on user feedback, allows an analyst to see more relevant information.

5 Conclusions

We developed CyberPetri, a prototype visualization tool based on previous work from Ocelot, for real-time network situation awareness at CDX 2016. Our tool showed the status of tokens and services across the subnets controlled by the participating service academies as these resources changed over time. Our tool organized and represented the event-based data through a containment metaphor using circle packing, and also provided supporting visualizations including a configurable timeline and horizontal stacked bar charts to provide additional context. From this experience, we elicited several additional requirements from the CDX participants that are leading to on going architectural and design improvements. Most importantly, these improvements will focus on **1) scaling to larger datasets** by aggregating many events into single nodes; **2) providing more exploratory capability** with the ability to focus on arbitrary time windows; and **3) improving situational awareness** through the ability to configure and then monitor more complex watches.

Acknowledgements

The research described in this document was sponsored by the U.S. Department of Defense (DOD) and by the U.S. Department of Energy (DOE) through PNNL. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government.

References

[1] D. L. Arendt, R. Burtner, D. M. Best, N. D. Bos, J. R. Gersh, C. D. Piatko, and C. L. Paul. Ocelot: user-centered design of a decision support visualization for network quarantine. In *Visualization for Cyber Security (VizSec), 2015 IEEE Symposium on*, pages 1–8. IEEE, 2015.

[2] T. Dwyer and L. Nachmanson. Fast edge-routing for large graphs. In *International Symposium on Graph Drawing*, pages 147–158. Springer, 2009.

[3] National Security Agency Information Assurance Directorate. Cyber defense exercise (CDX), February 2016. `https://www.iad.gov/iad/programs/cyber-defense-exercise/`. [Online; Last updated 1 June 2016].