ROYAL HOLLOWAY UNIVERSITY OF LONDON

Causality In Enterprise Systems

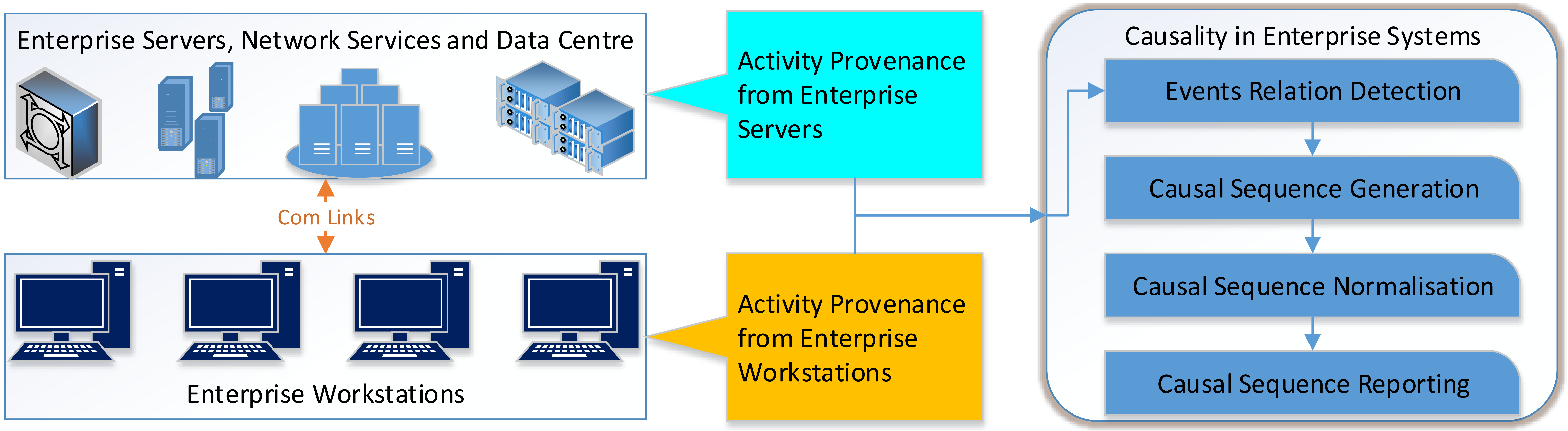
Freya Sheer Hardwick- Supervised by: Kostas Markantonakis, Raja Naeem Akram
Information Security Group, Smart Card and IoT Security Centre

The Smart Card and Internet of Things Security Centre

Objectives

The goal of the overall project is to create a system that enumerates logs from across multiple sources in an enterprise, creates a full causality chain of events, and presents these in a user friendly and easy to consume form. This goals of this section of the project include:

- To create a linking mechanism that connects isolated events together.
- To create a format for the event chains that is easy to search.
- To create a framework that evolves overtime to capture new events.



Enterprise Servers, Network Services and Data Centre

Com Links

Enterprise Workstations

Activity Provenance from Enterprise Servers

Activity Provenance from Enterprise Workstations

Causality in Enterprise Systems

Events Relation Detection

Causal Sequence Generation

Causal Sequence Normalisation

Causal Sequence Reporting

Figure 1: Data moving through a network

Introduction

Recent news pieces have illustrated the growing demand for transparency in how data is moved around and accessed in large enterprise systems. As it stands, the process to retrieve this infromation requires a laborious process of formal requests that result in incomplete and complex documents that are difficult for anyone, without the expertise or time, to understand. The purpose of this project is to create a framework that will automatically link together events occuring in an enterprise to track data as it is accessed, propegated, and altered. The end product will be a complete and thorough record of data movement. This is one puzzle picece in a larger project that aims to dismantle the barriers between an individual and the data that enterprises hold on them.

Initial Steps

The initial steps of this project involve making the connections between communicating entities. The main concern of our system is the actions taken by individual users, and to a lesser degree the devices that the users make the actions on.

The logs received by the framework will contain information regarding communications between IPs and the users that instigate certain actions.

This will be used in the formative stages to create a graph that models the landscape and heat spots of an enterprises communications. The data involved in these actions and communications will also be contained in these logs. Although not expressed in the graph, this data is the narrative that uses the graph as a foundation. In this way, the graph is like a story board with actors but no dialogue.

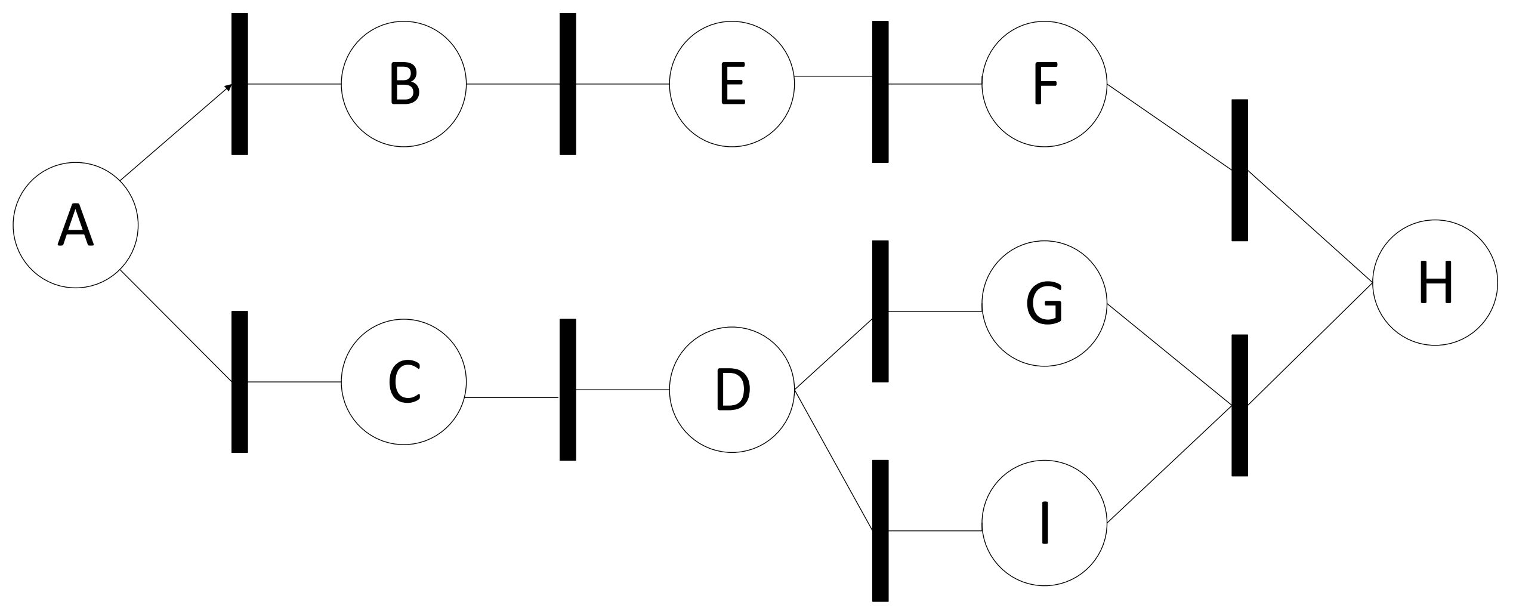


Figure 2: A petri net. One possible solution.

The matching of data to the graph nodes that concern it will allow us to maintain a full view of data movement. Due to the graph showing the actions of the users, rather than tracing the data, we will also be able to make connections between apparently disparate data accesses and data modifications. In this way, the framework moves away from simply capturing data in a different format to logs and provides its own value in the perspective it reveals.

Next Steps

The next stages of this project will be using the constructed graph to make inferences behind how data is being used. It's one thing to see huge swathes of data leaving the network but another to conclude that this data has been leaked or sold. We aim to give an individual full knowledge of how their data is being used.

Additional Information

This project is related to the EPSRC funded project "Data to Improve Customer Experience (DICE)". The project is particularly interested in personal data, and is using rail passengers as a specific focus of interest. The overall aims of the project are:

- Understand the role that personal data plays in enhancing the user experience of rail passengers
- To develop technical solutions to data privacy
- To develop an evaluation framework that can be implemented so passengers can understand how their data is used and how they can control and verify its use.

The project started in October 2016, and runs for three years to September 2019. For more information about the project, please visit http://www.dice-project.org

Acknowledgements

We acknowledge the support of the ISG-SCC for the summer internship program and EPSRC funded project. The views and opinions expressed in this poster are those of the authors and do not necessarily reflect the position of DICE project or any of partners associated with this project.

Contact Information

- Web: https://scc.rhul.ac.uk/
- Email: Freya.SheerHardwick.2016@live.rhul.ac.uk