Graph Databases for Use with Timeseries IoT Datasets

Colter Snyder
Department of Computer Science
Colorado School of Mines
Golden, Colorado, USA
csnyder1@mines.edu

***Abstract*—The wide spread proliferation of Internet of Things (IoT) devices have brought up many questions and concerns about their security and what they do with data. Many solutions have appeared using various solutions such as IoT Inspector [1]. However, these solutions don't make use of the unique capabilities and efficiencies that graph databases provide. This paper seeks to use a graph database in order to answer challenging questions about IoT devices particularly within the realm of timeseries datasets.**

***Index Terms*—IoT, graphical databases, databases, systems**

I. Introduction

Smart homes are everywhere now adays, from lights to door locks, TVs to speakers. It seems that these devices, which are refered to by their collective as Internet of Things or IoT devices, are in every home. With all these new devices come a whole slew of concerns about privacy and security [1]. There are many papers that explore these concerns, what this paper seeks to persue is how to efficiently analyze data collected from these devices such that people may infer various aspects about their devices. In particular, this paper seeks to see what what info can be garnered from timeseries datasets. Such aspects could include anything from what the device is doing to answering if a device is attacking the network and how.

It was decided that using a graph database could provide answers to these questions more efficiently and easily than a relational database. The primary advantage for use in this paper is the fact that graph databases are great for use with densely connected data [2]. On top of this, they are very quick to query and will give a result relatively quickly compared to other solutions [2].

II. Related Works

There are many current systems that implement different components of the general idea of graphical databases for IoT device management andanomoly detection, but none put these components together. *IoT Inspector* is a great tool that performs the task of monitoring using a form of a relational database, but not a graphical one [1]. The authors of *The Graph of Things* created such a system for aggregating IoT devices worldwide [3]; However, these systems are not built for small networks or home users.

III. Methodology

The first step in this project was choosing the technology that was to be used. It was decided to use Neo4J and Python for the tech stack.

IV. Results

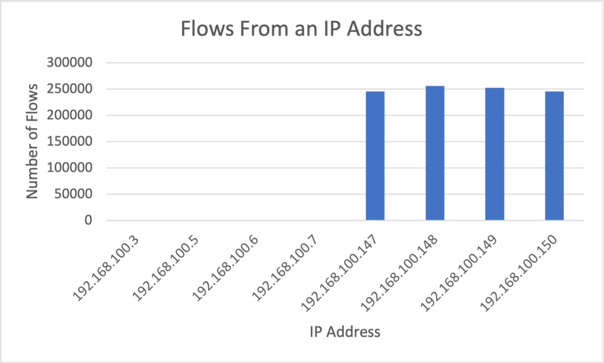

Fig. 1. The number of flows that were sent from particular IP addresses

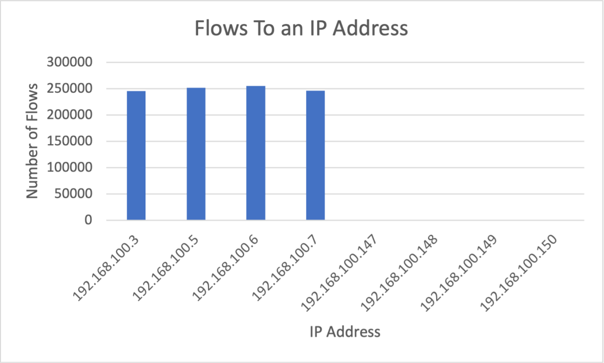

Fig. 2. The number of flows that were sent to particular IP addresses

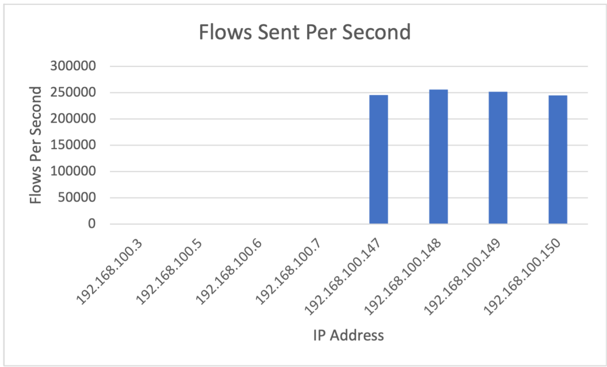

Fig. 3. The number of flows that were sent from particular IP addresses per second

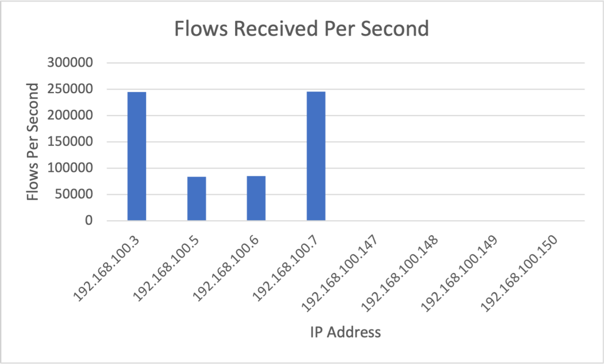

Fig. 4. The number of flows that were sent to particular IP addresses per second

V. Analysis

VI. Future Work

VII. Conclusion

References

[1] Danny Yuxing Huang et al. “IoT Inspector: Crowdsourcing Labeled Network Traffic from Smart Home Devices at Scale”. In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4.2 (June 2020). DOI: 10.1145/3397333. URL: https://doi.org/10.1145/3397333.

[2] Rohit kumar Kaliyar. “Graph databases: A survey”. In: *International Conference on Computing, Communication & Automation*. 2015, pp. 785–790. DOI: 10.1109/CCAA.2015.7148480.

[3] Danh Le-Phuoc et al. “The Graph of Things: A step towards the Live Knowledge Graph of connected things”. In: *Journal of Web Semantics* 37-38 (2016), pp. 25–35. ISSN: 1570-8268. DOI: https://doi.org/10.1016/j.websem.2016.02.003. URL: https://www.sciencedirect.com/science/article/pii/S1570826816000196.

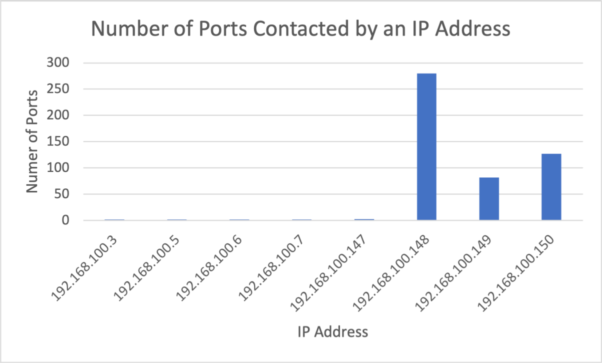

Fig. 5. The number of ports contacted per IP address

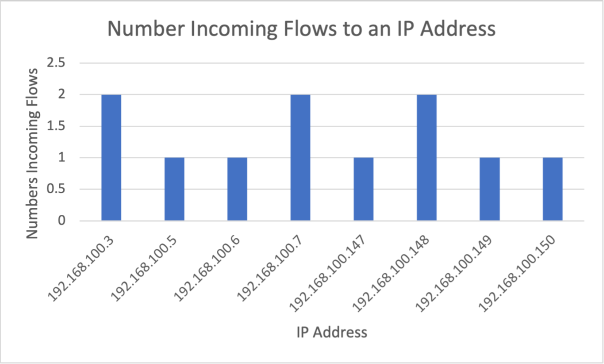

Fig. 6. The number of flows from unique IP addresses coming into a particular IP address