Hatrick: A System for Real-time Threat Detection in Cyber Physical Systems

Charith Wickramaarachchi, Alok Kumbhare
Computer Science Department
University of Southern California
Los Angeles, CA USA
email: {cwickram, kumbhare}@usc.edu

Charalampos Chelmis, Marc Frincu and Viktor K. Prasanna
Department of Electrical Engineering
University of Southern California
Los Angeles, CA USA
email: {chelmis, frincu, prasanna}@usc.edu

***Abstract*—Complexity of cyber attacks has grown rapidly over the last few decades. Novel advance techniques are needed in order to counter these attacks. Detecting some of these complex cyber attacks can be reduced to detecting patterns and dynamics in computer network traffic. These patterns can be molded as directed graphs based on their propagation through the cyber physical systems.**

This work in progress report presents an implemented system, Hatrick, which enable scalable, low latency dynamic graph analytics on clouds and commodity clusters. Hatrick will enable continuous monitoring of cyber physical systems to detect attack patterns in real-time.

I. Introduction

Cyber physical systems are becoming increasingly complex with the advancement of internet infrastructure, internet of things (IoT), and sensor networks. Smart grids and smart oil fields are an emerging class of such systems. In general, these systems consist of a large sensor network to maintain situational awareness.

Complex interactions happen between the interconnected cyber layer and the physical layer of these systems demands advanced analytics. These analytics should be able to correlate a large number of real-time data streams from distributed sensors, combine with historical data and detect events as they occur. Interactions in these systems generate high velocity data streams. Rich interrelationships between the data items in those streams can be modeled as graphs for analysis purposes.

Interaction patters in network accesses and online transactions can be used to detect cyber attacks [1]. Those patterns generally have repeating graph structures. Patterns in Witty Warm attacks, Smurf DDoS attacks and Fraggle DDoS attacks identified in the literature are some examples [1].

Smart grid is a key component when it comes to the realization of smart cities. It contains an advanced metering infrastructure (AMI) at customers and sub-stations, distributed sensor networks to maintain the situational awareness, communication network to communicate these sensor readings and distributed power generation, transmission and storage units. Real time state estimation in smart grid is one key problem when it comes to smart grid security. A dynamic graph can be created by modeling busses and branches of the electrical network as vertices and edges respectively. Tracking the connectedness of the network is used for state estimationand detect potential threats (ex: Creation of disconnected islands, bridges etc.) [2].

Detecting these dynamics as they happen is critical in order to minimize/avoid the damage that can be caused. This require development of low latency analytics. Large scale of the graph size, velocity of the data streams and low latency requirements of the analytics demands novel scalable systems to address these challenges. Existing large scale graph processing systems focus on batch processing of data which result in high latencies [3], [4]. Incremental algorithms has been proposed to address this challenge [5].

Hatrick enables incremental computation on large scale dynamic graphs. Hatrick leverages the elastic resource utilization capabilities available in the cloud to cater the low latency demands of applications. Hatrick performs efficient adaptive load balancing, auto-scaling and state management of the system with minimal data transfers between workers, while adapting to variable data stream rates. Programming model provided by Hatrick enables users to develop elastic incremental evolving graph analytics. To the best of our knowledge Hatrick is the first system to enable elastic dynamic graph analytics on cloud environments. Initial version of the system has been already demonstrated at CCGrid 2015 scale challenge[1].

References

[1] S. Choudhury, L. Holder, G. Chin, A. Ray, S. Beus, and J. Feo, "Streamworks: a system for dynamic graph search," in *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*. ACM, 2013, pp. 1101–1104.

[2] A. Abur and A. G. Exposito, *Power system state estimation: theory and implementation*. CRC Press, 2004.

[3] G. Malewicz, M. H. Austern, A. J. Bik, J. C. Dehnert, I. Horn, N. Leiser, and G. Czajkowski, "Pregel: a system for large-scale graph processing," in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*. ACM, 2010, pp. 135–146.

[4] Y. Simmhan, A. Kumbhare, C. Wickramaarachchi, S. Nagarkar, S. Ravi, C. Raghavendra, and V. Prasanna, "Goffish: A sub-graph centric framework for large-scale graph analytics," *arXiv preprint arXiv:1311.5949*, 2013.

[5] W. Fan, X. Wang, and Y. Wu, "Incremental graph pattern matching," *ACM Transactions on Database Systems (TODS)*, vol. 38, no. 3, p. 18, 2013.

[1] https://www.youtube.com/watch?v=uTDz5pQm8Bw