BIOGRAPHICAL SKETCH

Ramesh Paudel was born in Kaski, Nepal. He received a Bachelor's degree in Information Technology from Purbanchal University, Biratnagar, in 2011. He entered Tennessee Technological University in August 2012 and received a Master of Science degree in Computer Science in May 2014. In August of 2016, he was accepted into the Ph.D. program at Tennessee Tech University. He worked as a research assistant in the Data Science and Analytics Collaboratory under the supervision of Dr. William Eberle. He has published 7 peer-reviewed research articles in prestigious journal and conferences in the fields of graph mining, anomaly detection, and machine learning.

EDUCATION

Ph.D. Engineering
Tennessee Tech University, 2016-2020 (expected)

M.S. Computer Science
Tennessee Tech University, 2012-2014

B.S. Information Technology
Purbanchal University, 2007-2011.

Funding Acknowledgements

Funding provided by Tennessee Tech University, College of Engineering from the allocated “Carnegie funds”.

College of Engineering

TENNESSEE TECH

The Department of

Computer Science

Announces the Dissertation Defense

of

Ramesh Paudel

In Partial Fulfillment of the Requirements

For the degree of

Doctor of Philosophy in Engineering

April 2, 2020

3:00 pm

Held at

Tennessee Tech University

Johnson Hall Room 401

Zoom Link https://tntech.zoom.us/j/136784250

FIELD OF STUDY

Computer Science

DISSERTATION TOPIC

Efficient Graph Knowledge Discovery on Graph Streams with Concept Drift

EXAMINING COMMITTEE

Dr. William Eberle

Dr. Douglas Talbert

Dr. Gerald C. Gannod

Dr. Martha Kosa

Dr. Allan Mills

ABSTRACT

In a connected world like the world wide web, social networks, IoT sensor networks, computer networks, co-purchase networks, and telecommunications networks, where data represents the relationship or interaction between objects, graphs provide an effective representation. Yet, while machine learning applications can be used to make predictions or discover new patterns using graph-structured data, the complexity of graph data has imposed significant challenges. First, an efficient graph mining model needs to find a way to extract informative, discriminating, and independent features that can encode high-dimensional, interdependent, and irregular data in a format that can be used as input to a machine learning model. Second, with the increased use of mobile devices, internet-of-things technology, and online transactions, data are now more commonly available in a streaming fashion, thereby requiring the machine learning model to handle in one pass increasing graph volume arriving in real-time. Third, due to dynamic environments, changing internal/individual characteristics, and the intrinsic complexity of environments, the stream could undergo the phenomenon of concept drift (i.e., non-stationarity), leading to a deterioration of the model performance. Therefore, a real-time, graph learning model should be able to represent high dimensional data using informative and discriminative features, handle the streaming nature of data, and deal with concept drift. To address the aforementioned challenges of graph mining, we propose an integrated learning framework for a non-stationary graph stream that will i) handle the real-time graph stream in one-pass, ii) use a shingling-based sketching technique to represent a graph as a fixed-sized, informative, and discriminative feature vector, iii) identify concept drift using a discriminative subgraph-based approach, and iv) update the learning model after detecting concept drift so that the performance of the learning model does not deteriorate. Finally, we demonstrate the efficiency of the proposed framework using extensive experimentation on both synthetic and real-world graph streams.