University of Pittsburgh | School of Computing and Information

Dissertation Defense
Doctor of Philosophy in Information Science

"**Data reliability assessment based on Subjective Opinions**" by **Danchen Zhang**

Date: March 18, 2021
Time: 10:00am – 12:00pm
Place: https://pitt.co1.qualtrics.com/jfe/form/SV_1HBxzzPPYZHsbOe

Committee:

- Vladimir Zadorozhny, School of Computing and Information, University of Pittsburgh
- Daqing He, School of Computing and Information, University of Pittsburgh
- Vladimir Oleshchuk, Department of Information and Communication Technology, University of Agder
- Konstantinos Pelechrinis, School of Computing and Information, University of Pittsburgh

Abstract:
In the big data era, numerous data fluctuates society and people's life. These data come from diverse sources, and various information can be inferred and extracted. However, data quality usually cannot be guaranteed, and hence decision making with such unreliable data may lead to considerable losses. Accurate data reliability assessment mechanisms can help recognize the distrustful information and then filter unreliable data out.

In this work, I consider a novel approach to assess data reliability based on subjective opinions. I structure the data propagation model in terms of data sources producing and evaluating different statements. Next, I explore data history labels, value conflicts, and uncertainty. For different combinations of those parameters, I consider common scenarios, including handling fake news, truth discovery, data cleaning, as well as discovering cancer-driving genes.

In my dissertation, I explore how to accurately assess data reliability and how to make a decision based on evaluated reliability. I propose a series of subjective opinion based models to assess each scenario's reliability and compare them with state-of-art models through experiments on real-world data.