**Thesis Title:** Identifying misinformation and their sources in social networks


## ABSTRACT

Online social networks such as Twitter, Facebook and Sina Weibo are becoming increasingly important sources of information for its users. Misinformation such as rumors and fake news can spread quickly within social circles and may potentially incur massive losses to society. Therefore, it is often important to be able to accurately and promptly identify misinformation and their sources, so that proper control measures can be adopted. The research goal of this thesis is to develop methods estimating misinformation sources based on network topology and partial timestamps, and detecting misinformation based on surface-level linguistic patterns.

In many practical applications, the network topology may not be known in advance and needs to be inferred. We first consider the problem of inferring network topology from information cascades and knowledge of some moments of the diffusion distribution across each edge, without needing to know the distribution itself. We first propose an iterative tree inference algorithm and then generalized the algorithm heuristically to general graphs. Simulations using synthetic and real networks, and experiments using real-world data suggest that our proposed algorithm performs better than some current state-of-the-art network reconstruction methods.

We then study the problem of identifying misinformation sources in a network based on the network topology, and a subset of information timestamps. In the case of a single misinformation source in a tree network, we derive the maximum likelihood estimator of the source and the unknown diffusion parameters. We then introduce a new heuristic involving an optimization over a parametrized family of Gromov matrices to develop a single source estimation algorithm for general graphs. Compared with the breadth-first search tree heuristic commonly adopted in the literature, simulations demonstrate that our approach achieves better estimation accuracy than several other benchmark algorithms, even though these require more information like the diffusion parameters. We next develop a multiple sources estimation algorithm for general graphs, which first partitions the graph into source candidate clusters, and then applies our single source estimation algorithm to each cluster. We show that if the graph is a tree, then each source candidate cluster contains at least one source. We also propose a sequential source estimation algorithm using a particle filter that is based on an approximate hidden Markov chain model, which can be interpreted as a “reverse” propagation process. This algorithm allows the source estimate to be updated quickly whenever a new timestamp is available. Simulations using synthetic and real networks, and experiments using real-world data suggest that our proposed algorithms are able to estimate the true misinformation source(s) to within a small number of hops with a small portion of the information timestamps being observed.

Finally, we consider the problem of fake news detection based on surface-level linguistic patterns. We encode news text using a pre-trained and fine-tuned BERT language representation model. The full feature vector is obtained by integrating metadata with text. We develop a hybrid supervised learning framework combining a deep learning model and a graph model to predict news labels. We first train two deep neural networks, one to obtain the likelihood scores of different labels for any piece of news, the other to show the truthfulness relationship between any pair of news.  A graph is then constructed by adding edges between labeled and unlabeled news according to their predicted truthfulness relationships from the second neural network. For each unlabeled news, we compute another likelihood scores of different labels according to neighboring news with known labels. Combined with the likelihood scores given by the first neural network, we finally maximize the overall scores to obtain the label prediction. We apply our framework on the publicly LIAR dataset and obtained a 12% accuracy improvement over the state-of-the-art.

**PUBLICATION LIST**

## Journals

- *F. Ji, ***W. Tang**, W. P. Tay, and E. K. P. Chong, "Network topology inference using information cascades with limited statistical knowledge," Information and Inference, 2019, accepted. (*These authors contributed **equally** to this work)
- F. Ji, **W. Tang**, and W. P. Tay, "On the properties of Gromov matrices and their applications in network inference," IEEE Trans. Signal Process., vol. 67, no. 10, pp. 2624 – 2638, May 2019.
- **W. Tang**, F. Ji, and W. P. Tay, “Estimating infection sources in networks using partial timestamps,” IEEE Trans. Inf. Forensics Security, vol. 13, no. 2, pp. 3035 – 3049, Dec. 2018.

## Conferences

- F. Ji, **W. Tang**, and W. P. Tay, “Properties and applications of Gromov matrices in network inference,” in Proc. IEEE Workshop on Statistical Signal Processing, Jun. 2018
- **W. Tang**, F. Ji, and W. P. Tay, “Multiple sources identification in networks with partial timestamps,” in Proc. IEEE Global Conf. on Signal and Information Processing, Nov. 2017
- F. Ji, **W. Tang**, W. P. Tay, and E. K. P. Chong, “Inferring network topology from information cascades,” in Proc. IEEE Int. Symp. on Inform. Theory, Jun. 2017
- **W. Tang** and W. P. Tay, “A particle filter for sequential infection source estimation,” in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Mar. 2017