2014-2015 Data Mining Group Work Plan


Please note the following lists a set of sample templates. We expect you may discuss with your lab mates to dynamically form research groups to work on potential papers and/or systems. The product will be a strong research paper and/or an innovative research system prototype. Please do not forget to include (i) some interested, highly motivated master and bachelor students—note: Master and/or bachelor students may take lead to organize research groups as well, (2) visiting scholars and external collaborators (our ex-PhDs and some long-term collaborators, such as Nitesh Chawla (ND), Heng Ji (RPI), Jing Gao (Buffalo), Jessie Li (PSU), Yizhou Sun (NEU), Chi Wang (MSR), Xifeng Yan (UCSB) can be very good candidates for collaborations, (3) army researchers since we are required to work with them on several projects, including Lise Bowman, Hasan Cam, Jeff Hansberg, Tim Hanratty, Lance Kaplan, Sue Kase, and Clare Voss, and (4) other collaborators on research projects: Guohong Cao (PSU), Jim Hendler (RPI), Huan Liu (ASU), Al Wallace (RPI), Brooke Welles (NEU).

1. Network Construction
   - Joint entity typing (recognition) and summarization (clustering) from seeds
     - Xiang Ren, Urvashi Khandelwal
     - External collaborators: Heng Ji, Yizhou Sun
   - Event summarization by bootstrapping between discriminative phrases and time/location tags
     - Xiang Ren, Hyungsul Kim, Ahmed El-Kishky
   - Construction of large-scale hierarchical topic models supporting compounds
     - Jialu Liu
   - Concept network construction
     - Jialu Liu, Xiang Ren

2. Network Mining: Clustering, classification, prediction, evolution, and information diffusion analysis in heterogeneous information networks
   - ClusCite system: building effective citation recommendation system by information network-based clustering
     - Xiang Ren, Urvashi Khandelwal
   - Active learning in heterogeneous networks
     - Doris Xin, De Liao

3. Anomaly Mining: Mining anomalous sub-networks (e.g., terrorist ring) from massive, large networks (e.g., facebook) + outliers in networks [October monthly summary]

4. Truth Mining: Truth analysis among conflicting sources

5. CPS: Mining cyber-physical systems and networks

6. News Net Mining System: Research and development of a comprehensive newsnet network mining system