Semester summary

Yanglei Song
Overview

1. I spent some time working on the construction of HIN, mainly trying to use NLP tools to do preprocessing.

2. I mainly worked on the trust analysis. Next I’ll briefly introduce the recent progress in this field.
truth-finding vs. crowdsourcing

**Truth-finding:** very similar, but truth for each object can have different forms.

**Crowdsourcing:** outsource tasks to a distributed group of people. Tasks can be labeling images etc. (*Amazon Mechanical Turk*)
Previous work on crowdsourcing

I’m not saying crowdsourcing is well understood, but there are indeed many works on crowdsourcing which can be directed used in truth-finding.

NIPS09: model labeler ability, task difficulty

NIPS10: basically the same, but model each person may have different ability for different aspect
Recent work on truth-finding

Bo’s LTM

Latent Dirichlet Truth Discovery

Main contribution is to model the dependence of the sources, for example webs may copy content from other webs.

Bo’s main contribution is to model two-side accuracy of sources, allowing for multiple truth for each object.
1. Previous works have covered almost all factors we can think of, e.g. source quality, source dependence, difficulty of tasks etc.

2. But they only consider one factor at a time. So I’m thinking of integrating several factors in one model.