

# Extracting Relevant and Trustworthy Information from Microblogs

Krishna P. Gummadi

Max Planck Institute for Software Systems (MPI-SWS), Germany  
[gummadi@mpi-sws.org](mailto:gummadi@mpi-sws.org)

Microblogging sites like Twitter have emerged as a popular platform for exchanging real-time information on the Web. Twitter is used by hundreds of millions of users ranging from popular news organizations and celebrities to domain experts in fields like computer science and astrophysics and spammers. As a result, the quality of information posted in Twitter is highly variable and finding the users that are authoritative sources of relevant and trust-worthy information on specific topics (i.e., topical experts) is a key challenge. I will attempt to address this challenge in this two-part talk.

In the first part of the talk, I will focus on understanding and combating link farming activity in Twitter. Users, especially spammers, resort to link farming to acquire large numbers of follower links in the social network. Acquiring followers not only increases the size of a user's direct audience, but also contributes to the perceived influence of the user, which in turn impacts the ranking of the user's tweets by search engines. I will first discuss results from our recent studies investigating link farming activity in the Twitter network and then propose mechanisms to discourage the activity.

In the second part of the talk, I will focus on the problem of finding topic experts in Twitter. I will propose a new methodology that relies on the wisdom of the Twitter crowds. Specifically, we leverage Twitter Lists, which are often carefully created by individual users to include experts on topics that interest them and whose meta-data (List names and descriptions) provide valuable semantic cues to experts' domain of expertise. I will first describe how we mined List information to build Cognos, an expert search system for Twitter and then present results from a real-world deployment.

