# INFERRING PRIVACY INFORMATION FROM SOCIAL NETWORKS

Jianming He and Wesley W. Chu
Computer Science Department, UCLA

## Introduction:

Privacy confidentiality has become an important problem as online social network services grow in popularity. Based on the causal social relations among individuals, we study the property of inferring confidential information from these online networks [1].

## Methods:

Inferring whether an individual possesses a specific attribute from his/her social network can be a complex process. Using the fact that one's close friends tend to have more influence on that individual, we can reduce the size of these networks by making the following Localization Assumption: given the attribute value of X's friend at *i* (*i*≥1) hops away, Y, the attribute value of X is conditionally independent of the descendants of Y. To evaluate the inference probability on an individual via his/her friends, we shall leverage the Bayesian Network inference tools and map the social network onto a Bayesian Network. This requires us to convert the social network into a DAG by the following Naive Bayesian assumption: given the attribute value of X, the attribute values of Y (direct friends of X) are conditionally independent of each other. That is, the inference path from X to Y is considered as the primary correlation among these nodes, and the correlation among Y themselves can be disregarded as shown in Fig 1.

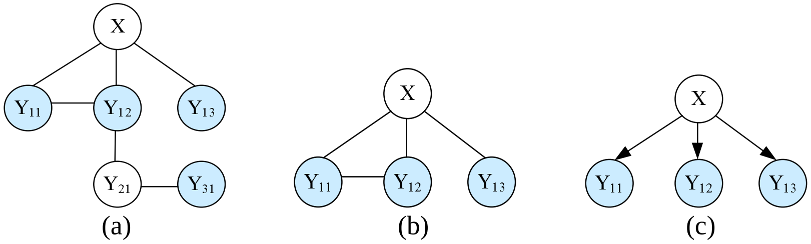

**Fig. 1** Reduction of a social network (a) into a Bayesian network to infer X from friends Y via Localization Assumption (b) and Naive Bayesian Assumption (c). The shaded nodes represent friends with known attributes.

## Results:

We study the impact of prior probability, influence strength and society openness on privacy inference. We crawl a social network of 66,766 people from Livejournal and compare the inference accuracy with social relations under different conditions. Experimental results reveal that privacy may be indirectly inferred via social relations, and the inference accuracy of confidential information is closely related to the inference strength between friends. Further, in a close society, privacy could still be inferred via Bayesian inference, despite the fact that people are hiding their attributes.

## Conclusions:

We studied the problem of inferring privacy information from social networks. Our results reveal that both the structure of the social network and their social relations affect privacy protection. We plan to investigate privacy protection by selectively hiding friendship relationships and/or requesting friends to hide their attributes under selected Social network structure and relations.

[1] J. He, W. W. Chu and Z. Liu. Inferring Privacy Information from Social Networks. To appear in Proceedings of *IEEE International Conference on Intelligence and Security Informatics*, 2006.