Privacy Wizards for Social Networking Sites

Shumin Guo

guo.18@wright.edu

September 15, 2010

Motivation

Why Social Networking Privacy?

- Online social networking privacy is one of the most urgent and implicit issues for online social network users.
- Users have difficulty reasoning about privacy and security policies, and thus difficult to make accurate privacy settings.

Then what can we do to aliviate the privacy threat for online network users?

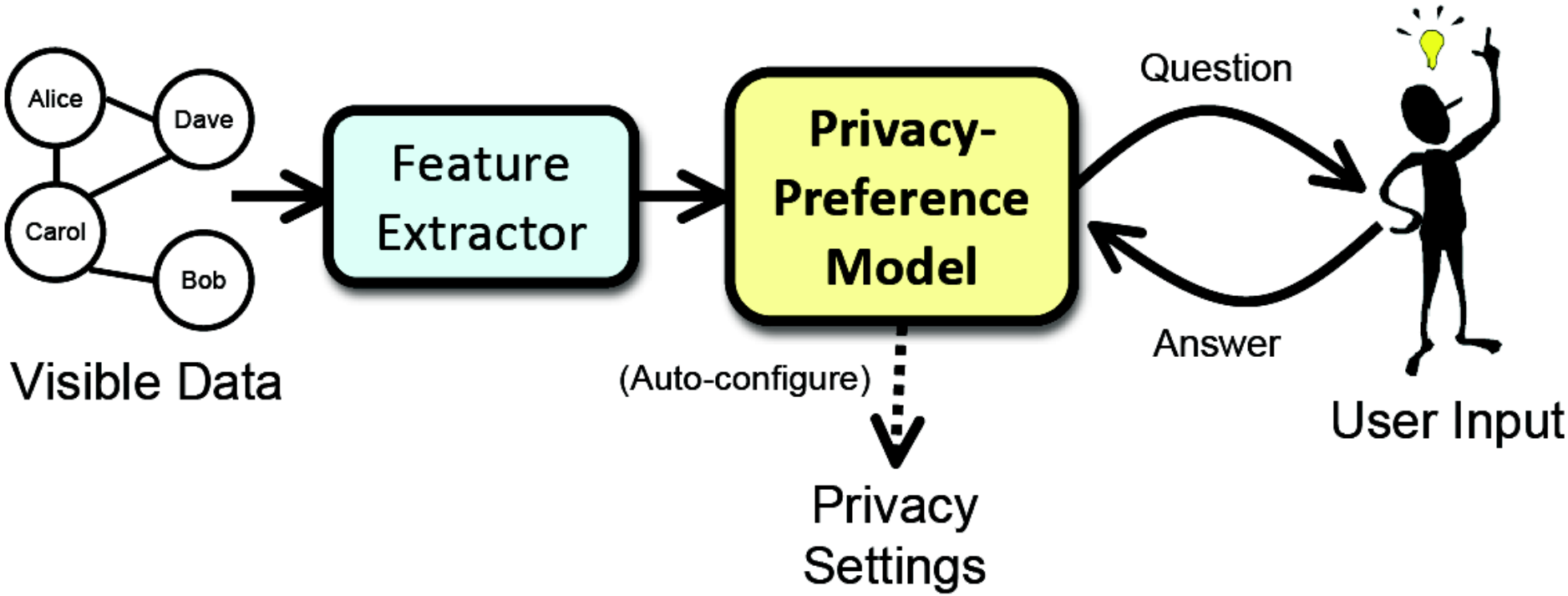
- Leave privacy settings to the default.
- Let users aware of the privacy threats and make more rational privacy settings.
- Automatically help users make privacy settings with a limited amount of user input.

Contributions of This Paper

- Build a machine learning model to describe a user's privacy preferences, with limited amount of user interaction(input); and incrementally optimize this model with more user input.
- Use the model to automatically configure user's privacy settings.
- Put user into a community environment, and recomment high-accuracy privacy settings using less user input than existing policy-specification tools.
- Visulization of the privacy wizard with a binary decision tree model.

Generic Privacy Model

With F being user's set of friends and I the set of information items in user's profile, we have $\widehat{pref} : IxF \rightarrow$allow,deny.



Wizard Specification

The privacy wizard solicits input from users by asking some simple questions as shown below:

Example (Example of wizard interaction with users)

```
Would you like to share DATE OF BIRTH with ...
Alice Adams? (Y/N)
Bob Baker?   (Y/N)
Carol Cooper?(Y/N)
...
```

According to the example above, it is natural to view the privacy-preference model as a binary classifier. This classifier can be infered from a set of labeled training examples (such as labeled friends) $F_{labeled}$. In this paper, the author uses feature vector $\vec{X}$ of a friend to predict the friend's privacy label. And the classifier can be viewed as a function of the form: $\widehat{pref} : \vec{X} \rightarrow$allow,deny. And the resulting classifier can be used to predict the user's privacy preference for unlabeled friends.

Feature Selection and Extraction

- Community Structure.
  We can extract a set of communities using $F_{labeled}$ and $F_{unlabeled}$.
  And then denote $G_1 = 1$ if user belongs to community G_1 and $G_1 = 0$ if not.
- Profile information of user and his friends.
  We can utilize the profiles, activities and other auxillary information(such as fans on facebook) of user's friends as feature
- Example of Extracted Features.

Name	Age	Gender	G0	G20	Fan_A	Label(DOB)
Alice Adams	60	F	0	0	1	allow
Bob Baker	16	M	0	0	1	deny
Carol Cooper	32	M	1	1	0	?

Uncertainty Sampling

This paper utilized Uncertainty Sampling to achive the best accuracy of the classifier with a limited amount of training data.

- Sampling phase
  The wizard selects friends for the user to label. The rule is to achive highest entrophy, as according to information theory, the higher the entropy, the more information the message will contain.
- Classifier Construction phase.
  In this phase, the wizard will use the labeled data to construct the classifier ($\widehat{pref}$), which is used to configure the user's settings.

Incremental Maintenance

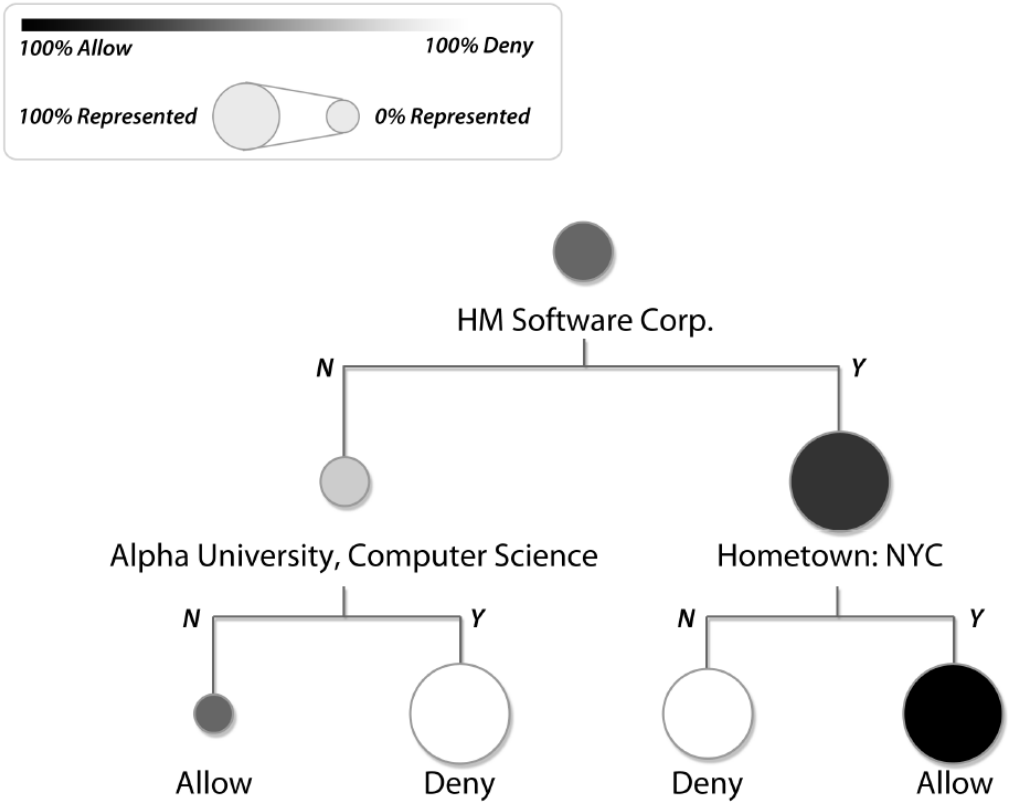
This paper considers the following two points concerned with incremental maintenance:

- Make reasonable prediction for new users without any additional input from user.
- In case user labeled new friends, these labeled data can be utilized in the classifier construction process.

With these two incremental maintenance considerations, the wizard can not only automate the prediction of user's preference on current friends but also the newly added friends and make the preference classifier more and more accurate with more input (labeled data/friends) from user. This will clearly satisfies the goal of this paper.

A binary decision tree visulization model

In this paper, the classifier wizard constructs a binary decision tree and present the result to the user a visulized tree model.



Evaluation

- Effectiveness.
- Most useful features for preference prediction.

The End