User Modeling and Recommendations

Many slides adapted from Lora Aroyo
http://de.slideshare.net/laroyo
Applications on the Social Web use web data [last week] & are ‘social’

To design ‘social’ functionality we need to understand how out of the data the application can provide relevant information (what users perceive as relevant)

Therefore we need to understand
- how good personalization (recommenders) are
- how good the user models are they are based on
- how to design and evaluate recommenders and user models (for use in SW applications)
Application has to obtain, understand and exploit information about the user

Information (need & context) about user

Inferring information about user & representing it so that it can be consumed by the application

Data relevant for inferring information about user
User Model

What may belong to a user model

- **Personal information**
  - Home address, telephone number
  - Friends information

- **Financial information**
  - Bank account, credit card

- **Work-related information**
  - Office address, office telephone number
  - Work colleagues

- **Preferences**
  - Topics
  - Purchases

- **Health Profile**

- **Citizen Profile**
  - Social security number
  - ID

...
People leave traces on the Web and on their computers:

- Usage data, e.g. query logs, click-through data
- Social data, e.g. tags, (micro-)blog posts, comments, bookmarks, friend connections
- Documents, e.g. pictures, videos
- Personal data, e.g. affiliations, locations
- Products, applications, services – bought, used, installed

Not only a user‘s behavior, but also interactions of other users

- „people can make statements about me“, „people who are similar to me can reveal information about me“ -> „social learning“ collaborative recommender systems
Austrian student takes on Facebook

By Sabrina Guillard (AFP) – Nov 12, 2011

VIENNA — Austrian law student Max Schrems may be just one of about 800 million Facebook users, but that hasn't stopped him tackling the US giant behind the social networking website over its privacy policy.

The 24-year-old wasn't sure what to expect when he requested Facebook provide him with a record of the personal data it holds on him, but he certainly wasn't ready for the 1,222 pages of information he received.

This included photos, messages and postings on his Facebook page dating back years, some of which he thought he had deleted, the times he had clicked "like" on an item, "pokes" of fellow users, and reams of other information.

http://www.google.com/hostednews/afp/article/ALeqM5gtxBJA6YlhCekdhyKKkZQhrRTk-g?docId=CNG.fa4291260bd695e9832eedb5a1266348.3f1
Reasons for protecting your social Web account
User Modeling Basic Concepts

- **User Profile**: a data structure that represents a characterization of a user at a particular moment of time represents what, from a given (system) perspective, there is to know about a user. The data in the profile can be explicitly given by the user or derived by the system.

- **User Model**: contains the definitions & rules for the interpretation of observations about the user and about the translation of that interpretation into the characteristics in a user profile.
  - *user model* is the recipe for obtaining and interpreting user profiles.

- **User Modeling**: the process of representing the user.
LET’S TALK ABOUT USER PROFILES FIRST
<table>
<thead>
<tr>
<th>Framework Type</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identity Framework</td>
<td>Authentication, Authorisation</td>
</tr>
<tr>
<td>Profile Framework</td>
<td>Distributed Attributes</td>
</tr>
<tr>
<td>Policy Framework</td>
<td>Privacy, Rights</td>
</tr>
<tr>
<td>Content Framework</td>
<td>Activity Streams, Syndication</td>
</tr>
<tr>
<td>Analytics Framework</td>
<td>Sentiment, Statistical</td>
</tr>
<tr>
<td>Other Frameworks</td>
<td>...</td>
</tr>
</tbody>
</table>

http://www.w3.org/2005/Incubator/socialweb/wiki/SocialWebFrameworks2
Identity Framework: Issues

- Let
  - Klout
  - Farmville
  - ...
- Use data from your Facebook account

- Naive approach
  - Klout/Farmville/... holds your FB login and password
  - Drawbacks:
    - Too many third parties gain access to your credentials
    - All-or-nothing access rights for Klout/Farmville....
    - Revocation is also all or nothing....
Oauth 2.0 (IETF)

- Authorization layer
- Separating roles
  - client (Klout/Farmville...) vs resource owner (you).
  - Resource server (Facebook)
- Client requests access to resources controlled by the resource owner and hosted by the resource server,
  - receives access token
    - more specific credentials than owner login/passwd
    - string denoting specific scope, lifetime & other access attribs
    - Are issued to third-party clients by an authorization server with the approval of the resource owner.
  - uses the access token to access the protected resources hosted by the resource server
Oauth 2.0 (IETF): Example

- **End-user (resource owner)**
  - can grant a printing service (client)
    - access to her protected photos stored
      - at a photo sharing service (resource server),
      - without sharing her username and password with the printing service.
  - Authenticates directly with a server trusted by the photo sharing service (authorization server),
    - which issues the printing service delegation-specific credentials (access token).
1. User tries to access your application.

2. Twitter then sends back the request token and request token secret.

3. After user allows the app, Twitter sends the Access token and Access token secret.

App goes to Twitter and gets the one-time request token.

App builds the authorization link, from which the user gets authenticated.

This token is unique for every user. App can store this in database for future access.

When app has the access token, it can access user’s details as permissible.

http://www.phpbuilder.com/columns/oauth-figure.jpg
User Profile

http://www.w3.org/2005/Incubator/socialweb/wiki/images/a/ad/1identity.png
Conclusion about Profile Framework

- Only considered in theory so far
- Closest match:
  - Centralized user profile: http://diasporaproject.org/
Policy Framework

- **P3P**
  - Platform for Privacy Preferences (P3P)
    - Domain-specific language
    - Web pages specifying what they do with your data (e.g. with cookies)
    - Your browser decides what to accept/reject
  - W3C recommendation since 2002
  - Mostly unused

- In general
  - Open research issue
  - Usability for developers and users remains a core open research issue!
Conclusion about Policy Framework

- Exists
- But is not used in practice
4 - Frameworks

- Identity Framework
  - Authentiction
  - Authorisation
  - ...

- Profile Framework
  - Distributed Attributes
  - ...

- Policy Framework
  - Privacy
  - Rights
  - ...

- Content Framework
  - Activity Streams
  - Syndication
  - ...

- Analytics Framework
  - Sentiment
  - Statistical
  - ...

- Other Frameworks
  - ...
  - ...
  - ...

http://www.w3.org/2005/Incubator/socialweb/wiki/SocialWebFrameworks2
Conclusion about Content Framework

- Crossposting
  - From twitter to facebook
  - From diaspora to facebook, twitter,....
USER MODELS
User Modeling Approaches

- **Overlay User Modeling**: describe user characteristics, e.g. “knowledge of a user”, “interests of a user” with respect to “ideal” characteristics

- **Customizing**: user explicitly provides & adjusts elements of the user profile

- **User model elicitation**: ask & observe the user; learn & improve user profile successively
  - “interactive user modeling”

- **Stereotyping**: stereotypical characteristics to describe a user

- **User Relevance Modeling**: learn/infer probabilities that a given item or concept is relevant for a user
Among the oldest user models

Used for modeling student knowledge

The user is typically characterized in terms of domain concepts & hypotheses of the user’s knowledge about these concepts in relation to an (ideal) expert’s knowledge

- Concept-value pairs
User Model Elicitation

- **Ask the user explicitly -> learn**
  - NLP, intelligent dialogues
  - Bayesian networks, Hidden Markov models

- **Observe the user -> learn**
  - Logs, machine learning
  - Clustering, classification, data mining

- **Interactive user modeling:**
  mixture of direct inputs of a user, observations and inferences
Hunch’s “Taste Graph” is a development platform for taste-based applications. Our self-serve API is free for non-commercial use.

Learn more about our API!

http://hunch.com
Stereotyping

- Set of characteristics (e.g. attribute-value pairs) that describe a group of users

- User is not assigned to a single stereotype – user profile can feature characteristics of several different stereotypes
Why are stereotypes useful?

http://farm1.staticflickr.com/155/413650229_31ef379b0b_b.jpg
Can we infer a Twitter-based user profile?*

* Example from Abel et al. (2011)
1. Which tweets of the user should be analyzed?

Profile:
- concept
- weight

(a) time period
(b) temporal patterns

1. Temporal Constraints

time

Morning:
Afternoon:
Night:

start
weekends
end

June 27
July 4
July 11
User Modeling Building Blocks

1. Temporal Constraints

2. Profile Type

Profile?

- hashtag-based
- entity-based
- topic-based

2. What type of concepts should represent “interests”?

Steffen Staab
staab@uni-koblenz.de
3. Further enrich the semantics of tweets?

Francesca Schiavone won! http://bit.ly/2f4t7a

Profile?
- Francesca Schiavone
- French Open
- Tennis

(a) tweet-based

Francesca wins French Open

(b) further enrichment

French Open
Tennis

Thirty in women’s tennis is primordially old, an age when agility and desire recedes as the...
4. How to weight the concepts?

Concept frequency (TF)
TFxIDF
Time-sensitive

Profile?

<table>
<thead>
<tr>
<th>concept</th>
<th>weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Francesca</td>
<td>4</td>
</tr>
<tr>
<td>Schiavone</td>
<td></td>
</tr>
<tr>
<td>French Open</td>
<td>3</td>
</tr>
<tr>
<td>Tennis</td>
<td>6</td>
</tr>
</tbody>
</table>

1. Temporal Constraints
2. Profile Type
3. Semantic Enrichment
4. Weighting Scheme

WeST
Steffen Staab
staab@uni-koblenz.de
Observations

- Profile characteristics
  - Semantic enrichment solves *sparsity problems*
  - *Profiles change over time*: fresh profiles reflect better current user demands
  - *Temporal patterns*: weekend profiles differ significantly from weekday profiles

- Impact on news recommendations
  - The more *fine-grained the concepts* the better the recommendation performance:
    entity-based > topic-based > hashtag-based
  - Semantic enrichment improves *recommendation quality*
  - *Time-sensitivity* (adapting to trends) improves performance
USER ADAPTATION

Knowing the user - this knowledge - can be applied to adapt a system or interface to the user to improve the system functionality and user experience.
Last.fm: Adapts to your music taste

user profile
interests in genres, artists, tags

user modeling
(infer current musical taste)

compare profile with possible next songs to play

history of songs, like, ban, pause, skip

next song to be played
Issues in User-Adaptive Systems

- Overfitting, “bubble effects”, loss of serendipity problem:
  - systems may adapt too strongly to the interests/behavior
  - e.g., an adaptive radio station may always play the same or very similar songs
  - We search for the right balance between *novelty* and *relevance* for the user (Diversity!)

- “Lost in Hyperspace” problem:
  - when adapting the navigation – i.e. the links on which users can click to find/access information
  - e.g., re-ordering/hiding of menu items may lead to confusion