A SEMANTIC RECOMMENDATION SYSTEM FOR MOBILE DEVICES: DESIGN AND DEVELOPMENT

UN RECOMMENDATION SYSTEM SEMANTICO PER DISPOSITIVI MOBILI: DESIGN E SVILUPPO

Relatore:
Chiar.mo Prof. Sonia Bergamaschi

Candidato:
Elena Parmiggiani

Correlatore:
Prof. Monica Divitini
Basit Ahmed Khan

NTNU

Det skapende universitet

Presentation outline:
1. Project information
2. Project characteristics
3. Research questions
4. Model of context
5. Recommendation system
6. Implementation
7. Results evaluation
8. Conclusions and future work

DBGroup
DATABASE GROUP

Thesis retrievable at: www.dbgroup.unimore.it

Preliminary information

- This work was carried on at **IDI** (*Department of Information and Computer Science*) at the ***Norwegian University of Science and Technology*** (**NTNU**), in Trondheim, Norway.

- It is part of the FABULA project (see *www.fabula.idi.ntnu.no*)

- Aim of FABULA is to find and develop innovative solutions to support (informal) collaborative mobile learning activities in a city-wide scenario.

- Users can learn from each other and from the environment where they are located→ *Situated learning* is the *outcome of the interaction with the social and the physical environment*.

- The FABULA system is grounded on the AGORA multi-agent framework.

Characteristics of FABULA

- ## Context awareness
  - According to *Dey*:

    "Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and the application themselves."

    Need to model of context and identify its dimensions.

- ## Environment as the protagonist of interaction.
  - According to *Rossitto*, an environment has:
  - ➤ A ***physical denotation***, i.e. its geographical description: from this perspective, we call it a ***space***;
  - ➤ A ***connotation*** changing from user to user and resulting from his activities, his goals, his interests and his past experiences: in this case, we call it a ***place***;
    - → a place is the superstructure built ad-hoc for one user out of one common geographical space, using his explicit and implicit information.

Research questions

The two main goals of my work are:

a) To exploit ontologies and semantic technologies to obtain a **model of context** for a mobile learning application;

b) To justify and develop a set of semantic-based **recommendation services** to suggest relevant content to the user.

Meta-model of context

- The context modeling process was carried on in two steps:

  1. Development of a meta-model by means of a class diagram;

  2. Translation of the meta-model into an ontology.

- Categories (dimensions) of context:
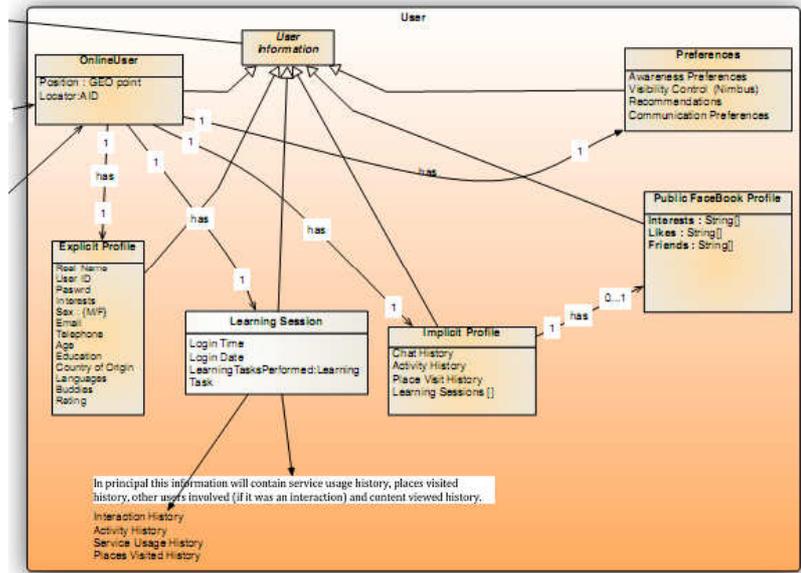
  ❖*Users*;
  ❖*Places*;
  ❖*Activities*.

The *users* dimension

Explicit information:

- *Learning session;*
- *Geographical position;*
- *Explicit FABULA profile;*
- *General visibility and awareness preferences.*

Implicit information:

- *Activities done history;*
- *Places visited history;*
- *Chat logs;*
- *Facebook profile.*

The *places* dimension

Division of the space into:

- **City (Trondheim);**
  - **Learning places;**
    - **Learning opportunities.**

Each of these levels has a *profile with*:

- *A set of tags;*
- *A description in free text;*
- *A meaning, represented as the intersection of three different layers:*
  - *Historical,*
  - *Social,*
  - *Cultural.*
- *A 1-10 rating value for each level.*
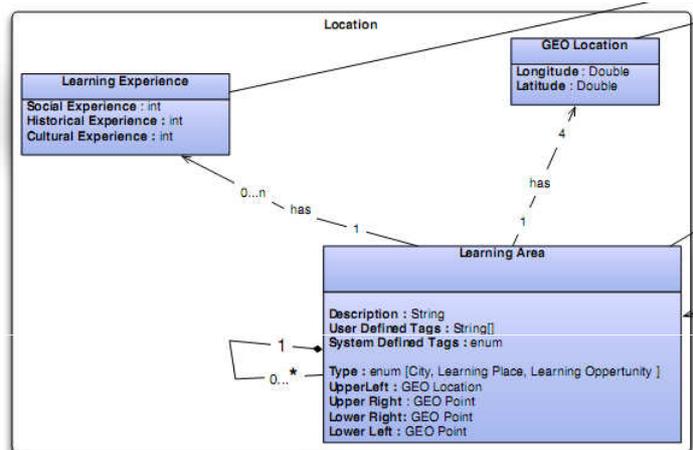
The *activities* dimension

- **An activity (solo-based or group-based) is a sequence of tasks that the user must accomplish and takes place in a given learning place;**

- **The structure of the *profile* used to describe an activity is formally similar to that of a place:**

  - *A set of tags;*
  - *A description in free text;*
  - *A meaning, represented as the intersection of three different layers:*
    - *Historical,*
    - *Social,*
    - *Cultural.*
  - *A 1-10 rating value for each level.*

The translation into an ontology

➢ *Representation:*

→ Context is a kind of dynamically changing knowledge: need to foster *dynamicity*, *reasoning* and *platform-independence*.

→ Solution: **OWL-DL ontology** developed with Protégé 4.1, divided in two levels:

1) A *general, upper ontology*;

2) A set of *domain-specific second-level ontologies*.
    → so far, only for the user dimension a specific ontology was realized by extending the **FOAF** vocabulary (see http://www.foaf-project.org/).
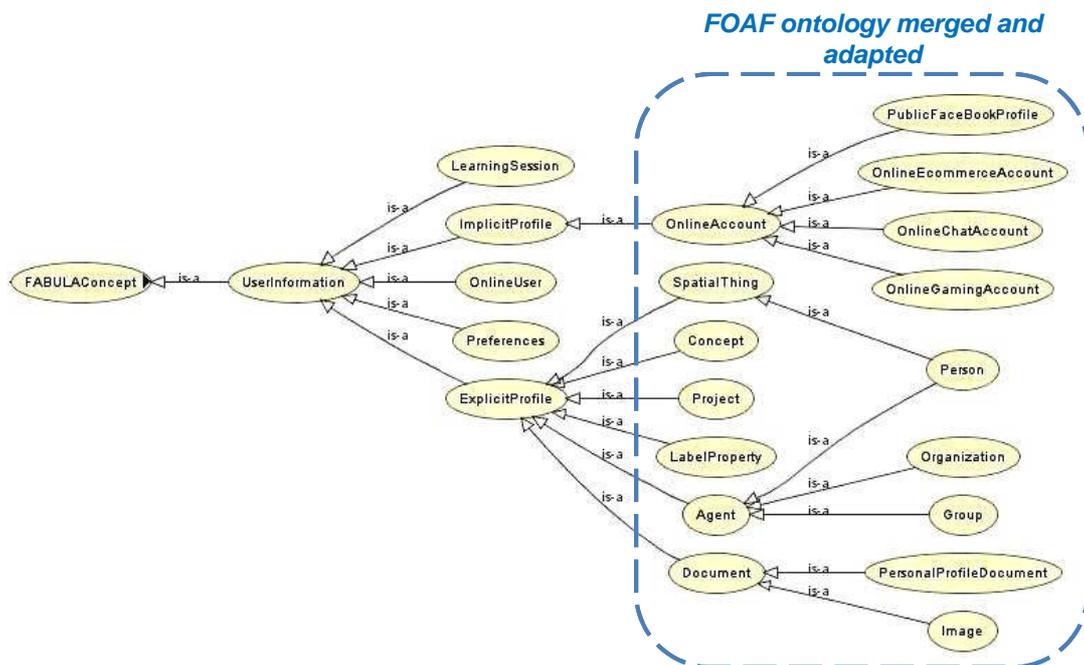
➢ *Modeling approach:*

→FABULA is built upon the **AGORA** multi-agent middleware → the structure of contextual knowledge has to respect the behavioral architecture followed by agents.

→ **CommonKADS** methodology (see http://www.commonkads.uva.nl/).

The upper FABULA ontology…
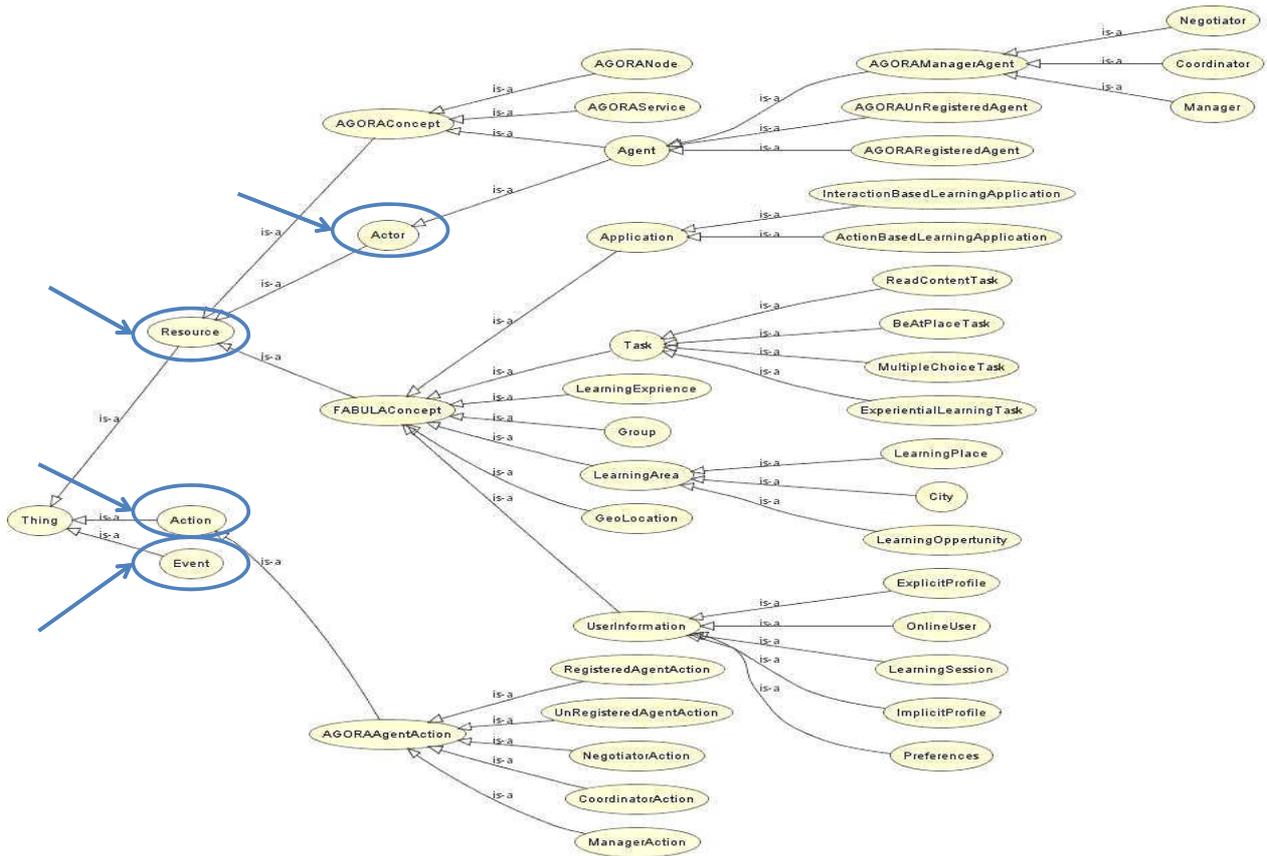
… and the user-domain ontology

The recommendation system (RS)

✓ In order to present relevant content belonging to one of the dimensions to a user, I wanted to gather the positive aspects of the RS available in literature:

– **Collaborative filtering** (opinion of similar users)
– **Content based** (comparison between user and item profiles)

An algorithm was designed to recommend the user items of one of the dimensions of context:

– A user (or a *friend*);
This part was implemented and the results of the techniques adopted where evaluated;
– A learning place or opportunity to visit;
– An activity to get engaged in.

Recommending a friend (1/2)

- The recommendation value $r\ (u_a, u_b)$ of user u_a to user u_b was achieved by matching the information about u_a and u_b:

$$r\ (u_a, u_b) = implicit\_profile\_matching\ (u_a, u_b) + explicit\_profile\_matching\ (u_a, u_b)$$

- Two matching phases:

  ➢ *Keyword-based matching* → be items in (interests, likes, languages spoken, friends,…)
  $$matching\ (u_a, u_b) =\ (items(u_a) + items(u_b))\ /\ (items(u_a) \cap items\ (u_b))$$

  + *Semantic matching* → *semantic distance* between words in the `is-a` hierarchy of *WordNet*.
    ▪ *path* measure: this solution does not rely on corpora-based probabilities, as we cannot make any assumption about the meaning of the concepts retrieved:
 $$sim(w_1, w_2) = 1/shortest\_path(w_1, w_2)$$

    ▪ **Similarity threshold**: $T = avg\_sim\_value + k * st\_dev,$
    (where k=1 for results evaluation).

Recommending a friend (2/2)

→ *Trustworthiness* value (or, at opposite, a *penalization*) for a user excessively active (so his matching value with another is less relevant) or too little active (he did not provide interesting info), if compared to the average behavior of all the FABULA users with a Facebook profile:

  ➢ The **average behavior** is the average nr of items (interests listed, items tagged as *I like it*, friends);

  ➢ five **levels of penalization**, directly proportional to the standard deviation.
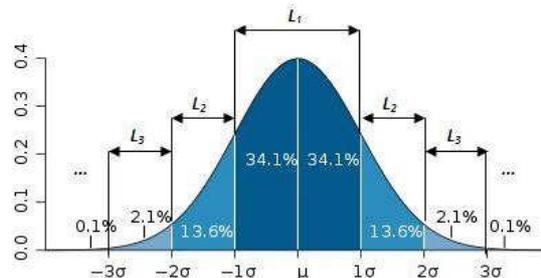

Recommending an item

- A common procedure was developed for recommending a ***place*** or an ***activity*** (from here on, *item*) to a user, as the structure of their profiles is overlapping.
- Techniques from Information Retrieval are exploited.

- Two approaches based on the data available have been designed:

  ➢ **Cluster-based matching** (content-based): items are clusterized according to their cosine similarity.
    - Stop condition to the clustering phase: at least one cluster has to reach a number of elements of ***1/heterogeneity***, where **heterogeneity** is the number of the different types of items available;
    - The more items from the past history of a user are present in a cluster, the more the other items inside that cluster are weighted with the **hyperbolic tangent** in the interval [0,1].

  ➢ **Similarity-based matching** (content-based + collaborative filtering):
    - Cosine similarity item-user;
    - Weight the similarity on the opinion (vote) of the users that are most correlated .
      → definition of a ***neighborhood***.
    - Do the same for every level of interest (historical, cultural, social).

Recommending an item: schema
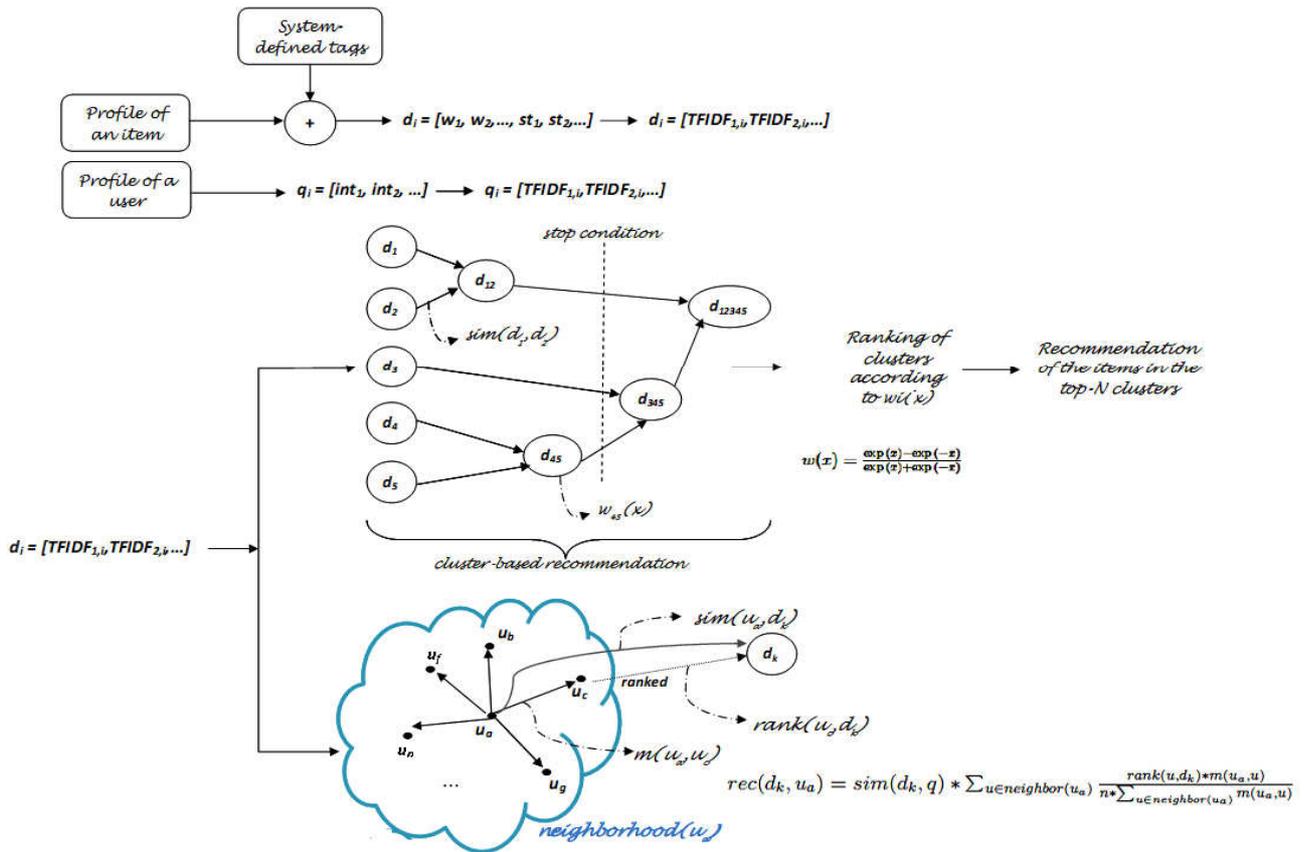
Recommending a friend: implementation

- The algorithm was implemented in *Java*.
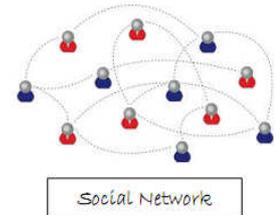
→ Interface to the context ontology to retrieve this information by means of the framework *Jena*.

→ Interface with Facebook by adapting the *Graph APIs.*

→ Interface with WordNet: **JWNL** (*JavaWordNetLibrary*).

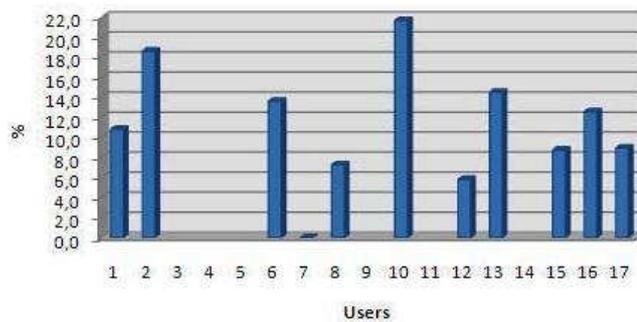# Results evaluation (1/2)

Dataset:

- 17 FABULA user profiles

- 11 Facebook profiles

- Around 10 activity profiles and 20 places and opportunities profiles.

➡️ *Results of the test phase for the recommendation of a friend:*

➢ Adding the information from Facebook, when available, improves the average correlation between one user and the others up to 11%, with peaks of 41%.
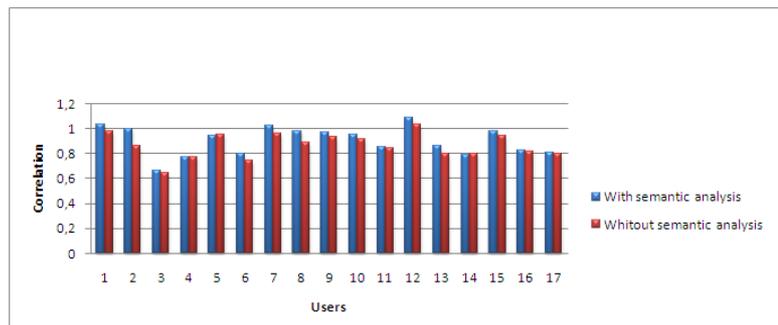

Results evaluation (2/2)

➢ Adding a penalization procedure (or trustworthiness assignment) to the Facebook users reduces the average correlation between a user and all others of around 2%;
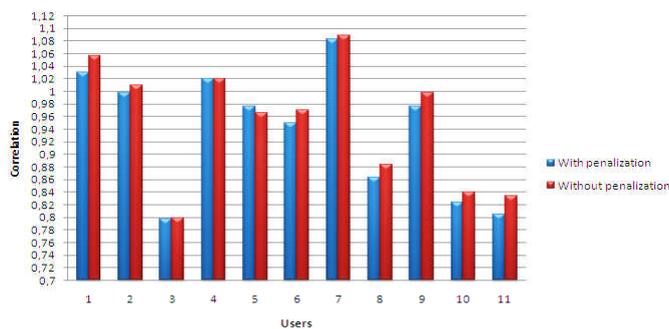
➢ Adding a matching procedure based on the meaning of concepts improves the average correlation of 7%.

Conclusions and future work

- Looking for a **specific ontology** to better describe a **place** by going into the theory about place and space.

- For the correlation between users:

  - ➢ Testing of **different semantic distance measure**, maybe based on the information content of a concept;

  - ➢ Testing of interfaces with **other Social Networks**, to also search for different types of items (music,…);

  - ➢ Evaluation of **different penalization algorithms** based on different paradigms (e.g., see Clarke Tax algorithm).

- Implementation of the whole algorithm developed here into a *Recommender Agent* inside the MAS of FABULA.

Thanks for your attention