

Hoarding Content in an M-Learning System

Anna Trifonova and Marco Ronchetti
University of Trento, Italy

outline

- Mobile learning
- What is hoarding and why it is needed?
- Our scenario
- The hoarding process
 - Learning sequences
 - Hoarding starting step
 - Analyzing user behavior
- Hoarding - will it work?
- Conclusions and future work

mobile learning



- E-learning + mobile computing
 - PDA
 - Cell phones
 - ❖ Any small, autonomous and unobtrusive device
- M-learning can include anything from job aids and courseware downloaded on your personal digital assistant to Net-based, instructor-facilitated training via laptop
- Interact with people
 - Voice
 - Written messages
 - Still and moving images
- Access content
 - Local
 - Remote
- Access services



hoarding– what is it? why we need it?

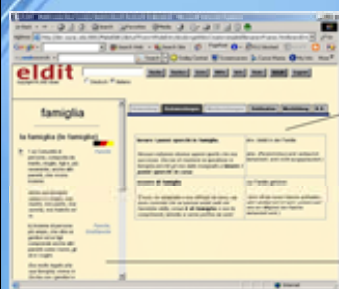
- Hoarding is a technique for selecting set of documents to be uploaded and used when disconnected.
- To allow 'any time, anywhere' access to the learning materials
 - to support offline access of learning content
- Often the device available memory is not big enough to contain all material of a system
 - a decision should be made what to put on the device
- To free the user from annoying procedures of pre-fetching content
 - decide automatically what the user will need

our scenario

- E-learning platform that containing big quantity of learning materials
- A self-motivated learner, who doesn't need any supervisory control of studying process
- A mobile device (PDA) which has intermittent connection and can not hold all the data of the e-platform in the memory
- The device is regularly synchronized with the main system

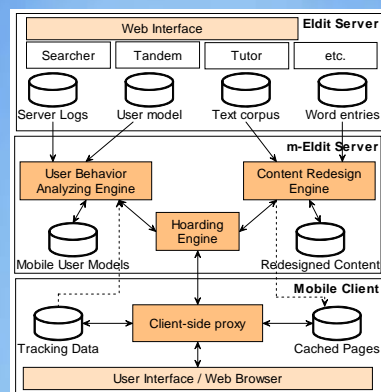


our scenario – more concrete



- **Content Redesign** – XML to HTML with XSLT
- **Client-side Proxy** – to respond to client requests offline and to keep tracking data
- **Hoarding and User behavior analyzing** engines on the server side – to analyze the tracking data; to create user models and to decide on what materials should be included in the hoarding set

- ELDIT -An innovative electronic language learning system that helps to prepare for the exams in bilingualism in South Tyrol and for everybody interested in learning German or Italian language
- ELDIT consists of: electronic learner's dictionary, a text corpus, quizzes, a tandem module, an adaptive tutor, etc. (about 800 texts and highly interconnected associated word entries) – **More than 600MB data**
- **MOBILE ELDIT** – for PDAs (**32-64MB memory**)



the hoarding process

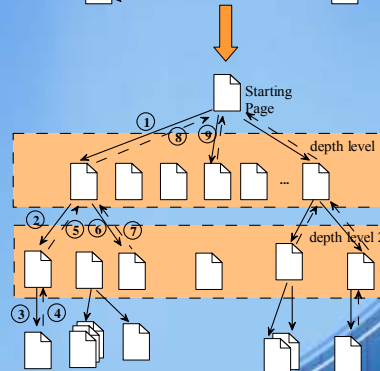
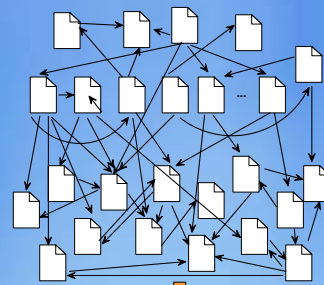
1. Decide (predict) what material the learner will start from;
2. From the materials that are linked to this 'start point' generate a 'candidate set';
3. Predict which will be the learning sequence or set, that is of interest for the user at next depth level and discard the items that are not of interest from the set (this step is called pruning) – loop until the deepest level;
4. Set priorities to the materials in the hoarding set, based on their predicted priority of importance for the next session (higher if the probability the object will be used soon is high);
5. Pre-fetch, starting from the beginning of the ordered list, putting on the device those objects with bigger priority, until available memory is filled in.

hoarding - questions

- What is the best starting point for the user's next session?
- What is a 'session' in the mobile learning scenario?
- How can we predict the most probable learning path (sequence of LO)?
- How do we create (formalize) a useful user model?
- What are the important parameters of the user behavior which have influence on the prediction?
- How do we use different parameters of the user model for predicting and/or pruning and do these different parameters have different significance for the prediction and/or pruning process?
- How do we formalize the domain knowledge?
- Do different domain knowledge parameters have different significance for the prediction and/or pruning?
- How do we measure the successfulness of the automatic hoarding and how do we improve the work of the algorithms, considering these measures?
- Etc.

hoarding - learning sequences

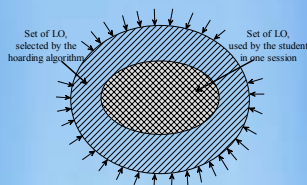
- Web-based materials are generally highly interlinked – a net of links
- The user browsing path over a web-based material can be viewed as a hierarchy structure
- Depends on user's learning style, natural learning habits and abilities



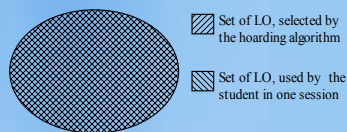
USER ANALYZING AND MODELING

hoarding - the starting step

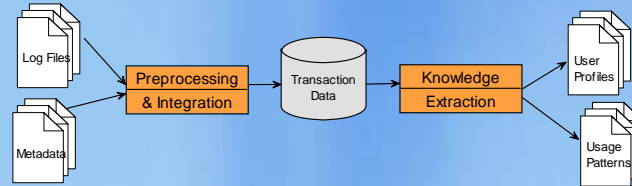
- 'Starting Point' → from site structure
- Based on observations on all previous users the system can estimate the average (or max.) browsing depth and session length



- Maximize the hit rate
Number of hits divided on number of uploaded items
- Minimize the miss rate
percentage of ineffective requests



analyzing users' behavior /1



- Generic users
 - Typical usage patterns
 - Categorization of users
 - Relations between items
 - etc.
- Particular user
 - Preference
 - Learning style
 - Knowledge obtained
 - etc.

analyzing users' behavior /2

- Association rules, like
 $LO1 \Rightarrow LO6 ; LO2 \Rightarrow LO1 ; LO2 \Rightarrow LO3 ; \dots$

"Almost every time when the LO1 was viewed also LO6 was viewed in the same session"
 (where LO1 might be an example problem and LO6 the solution given by the lecturer)

- Clustering

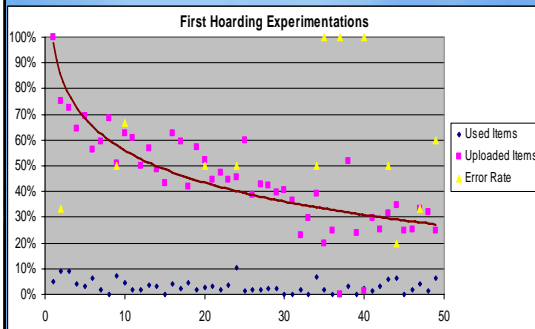
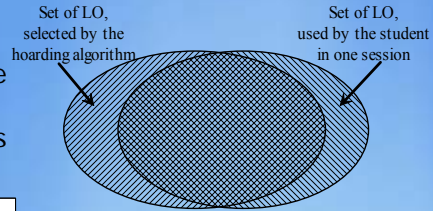
Cluster	Instances	Centroids	Associations
cluster 0	4 (57%) Session ₁ ; Session ₃ Session ₄ ; Session ₇	0 0 0 0 1 1	LO ₁ ⇒ LO ₅ LO ₃ ⇒ LO ₅ LO ₃ ⇒ LO ₆ LO ₄ ⇒ LO ₆
cluster 1	3 (43%) Session ₂ ; Session ₅ Session ₆	1 1 1 0 0 0	LO ₁ ⇒ LO ₃ LO ₃ ⇒ LO ₁

- Collaborative Filtering

$$\text{similarity}(\vec{A}, \vec{B}) = \cos(\vec{A}, \vec{B}) = \frac{\vec{A} \cdot \vec{B}}{\|\vec{A}\| * \|\vec{B}\|}$$

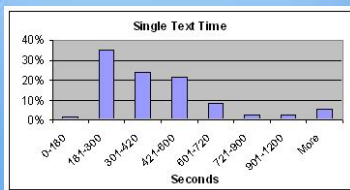
hoarding - will it work?

- Initial experiments of M-Eldit
 - Observe one user at a time
 - Do pruning by subtracting what the user knows
 - What the user decided not to see is what the user knows

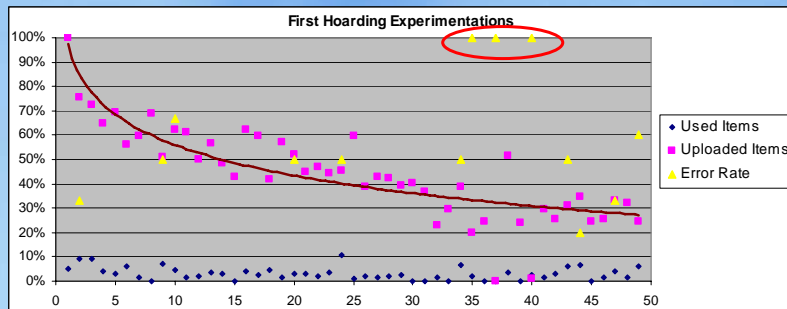


- Faster - combining the data of 2,3,... users
- More precise - by introducing 'Critical Set'; keep an eye on usage times, etc.

hoarding - usage times



Adding some 'intelligence', like taking different decisions depending on the usage times will lead to **higher precision** (i.e. lower error-rate)



conclusions & future work

- Hoarding is important, because we want to support learners on the move without adding the need of manual selecting and uploading of study content
- We outlined a general hoarding algorithm and sketched some possible techniques and approaches. We still have to deeply study the parameters that have influence on the algorithm efficiency and their weights
 - Effective pruning
 - Tuning of the algorithm
 - Which combination to use
- How can we generalize a strategy, so that the hoarding algorithm can be used in different mobile learning systems

Hoarding Content in an M-Learning System

THANK YOU!

Anna Trifonova
Trifonova@science.unitn.it

Marco Ronchetti
Ronchetti@dit.unitn.it

University of Trento
Via Sommarive 14, Povo (TN), ITALY

<http://www.science.unitn.it/~foxy/>