ERCIM-98-W001
SZTAKI

European Research Consortium
for Informatics and Mathematics

ERCIM

Fifth DELOS Workshop

Filtering and Collaborative Filtering

Budapest, 10-12 November 1997

FIFTH DELOS WORKSHOP

Filtering and Collaborative Filtering

Budapest

10-12 November 1997

CONTENTS

INTRODUCTION

The Fifth Workshop of the DELOS working group, held in Budapest, 10-12 November, 1997 organized by MTA SZTAKI, Department of Distributed Systems, gathered almost 30 experts from Europe and the United States to examine and discuss issues related to collaborative filtering techniques on the web and digital libraries.

The DELOS Working Group, part of the ERCIM Digital Library Initiative, is funded by the ESPRIT Long Term Research Programme (LTR No. 21057) within the Fourth Framework Programme of the Commission of the European Union. Its objective is to promote research into the further development of digital library technologies, in particular to (i) stimulate research activities in areas which are relevant for the efficient and cost-effective development of digital library systems, (ii) encourage collaboration between research teams working in the field of digital libraries, and (iii) establish links with on-going projects and activities in the field of digital libraries in industry and other public and private institutions.

Collaborating filtering (social filtering) systems aim at automating the "word of mouth". Relying on recommendations given by others usually happens in situations with either too much or too few information available. This is the current situation within Internet. Internet provides large quantity of information with different quality. Low quality information (junk) prevents users to find relevant information in an effective way. Mimic the most successful available social filtering systems, such as the job of journal/conference editors, or traditional recommendation systems for movies/books/CDs/etc., new information technologies, the collaborative filtering techniques are emerging. Collaborative filtering techniques are based on the collection of user ratings on information entities of the net. Users can help each other to distinguish between the high and low quality entities providing their opinions (in the form of ratings). Users may be guided by these ratings. Collaborative filtering systems collect these individual ratings and present them in an organized way. Global information infrastructure is necessary for the implementation of this idea. The Fifth DELOS Workshop targeted this R&D area, particularly the following topics of the workshop were planned:

- information filtering methods and algorithms
- rating techniques
- intelligent filtering and rating
- social (collaborative) rating and filtering
- agent-based filtering
- user modelling
- quality control of digital documents
- software tools for information filtering and rating
- searching methods and techniques
- resource indexing.

During the workshop four invited speakers presented their views on key issues in the field:

- Joe Konstan from the University of Minnesota gave a presentation of the GroupLens Research Project: Scalable Collaborative Filtering for the Internet.
- Jacob Palme from the Stockholm University and KTH discussed 'an architecture for intelligent and collaborative filtering'.
- Damian Arregui and Manfred Dardanne from the Xerox Research Centre Europe gave an overview of 'Knowledge Pump: Community-centered Collaborative Filtering'.

The presentations from European researchers covered several EU projects using collaborative filtering techniques such as EUROgatherer and SELECT of the Telematics Application Development Programme, a personalized information gathering system; SOAP, Live Recommendations through Social Agents; the TREVI Project - Personalized Information Filtering, Linking and Delivery for the News Domain; as well as novel approaches and techniques related to collaborative filtering: a software prototype for information filtering and rating using evolutionary algorithms; a direct manipulative tool for assembling profiles; the Use of LDAP in a filtering service for a digital library; usage, rating & filtering; institutional rating in everyday life; the application of a generic voting tool for rating purposes; social filtering and social reality; lightweight collaborations for social filtering on the web; a language theoretical approach to filtering and cooperation; a non-monetarian collaborative cooperation model in an Internet based groupware service; a visual tagging technique for annotating large- volume multimedia databases and a tool for adding semantic value to improve information rating.

A panel session with the invited speakers gave the participants an opportunity for a lively discussion on the topics raised during the workshop.

We thank the participants of the workshop for their presentations, and for the vivid interesting discussions during the workshop. We also thank ESPRIT and ERCIM for their contribution to the organization of the workshop and the publication of these proceedings.

László Kovács
MTA SZTAKI, Department of Distributed Systems

Project Overview: EUROgatherer - A Personalized Information Gathering System

Umberto Straccia
I.E.I. - C.N.R.
Pisa, Italy
http://faure.iei.pi.cnr.it/~straccia

1 The EUROgatherer project

The aim of this short note is tho describe the EUROgatherer project. The project is a 20 month project of the european Telematics Programme and will start in January, 1998, involving the following partners:

- I.E.I. - C.N.R., Pisa - ITALY, **Coordinator**
- Italia Online SpA, Milan - ITALY
- Rank Xerox Research Center, Grenoble - FRANCE
- Eurospider Information Technology AG, Zurich - SWITZERLAND
- Xarxa CINET SL, Barcelona - SPAIN
- University of Dortmund, Dortmund - GERMANY
- Dublin City University, Dublin - IRELAND

1.1 Rational of the project

A tremendous amount of news and information is created and delivered over electronic media. This has made it increasingly difficult for individuals to control and effectively manage the potentially infinite flow of information. Ironically, just as more and more users are getting on-line, it is getting increasingly difficult to find information unless one knows exactly where to get it from and how to get it. Tools to regulate the flow are urgently needed to prevent computer users from being drowned by the flood of incoming information. Traditional information retrieval systems concentrate on retrieval of unstructured texts of static documents. Information filtering systems have instead been applied to document streams, such as newswire, news groups, and electronic mail. Information gathering is a new field which combines features from information retrieval, information filtering, natural language and knowledge representation, and applies it to the new domain of documents structured in various forms (hypertext, MIME, etc.) and different formats (text, PostScript, GIF, MPEG, etc.). This field has recently seen a significant growth and an enormous popularity with the appearance of several search engines, such as Altavista, Lycos, Yahoo, Excite, Harvest, which help in finding material on the Web. These systems regularly scan the Web to produce indexes to be used in answering queries from users. They provide a generalized service of indexing digital collections accessible through the Internet. In essence, they index textual documents, structured in HTML pages, and provide a search and retrieval service to the Web community at large, but do not provide any personalized support to individual users. Indeed, they are targeted towards a general and generic user, and therefore they are oriented to answering queries crudely rather than to learning the long-term requirements idiosyncratic to a specific user and selecting and organizing material for him/her accordingly. The technology of information gathering can be applied to a huge number of on-line services, assisting for instance in the selection of books or other archived documents from libraries, news items from press agencies, television station and journals, or documents from administrative bodies. The niche for personalized, prioritized information as an alternative to the uniform newspaper or television broadcast media available today is likely to be the first application domain in which personalized information gathering systems become widespread.

The EUROgatherer project aims at designing and implementing a system which provides a personalized information gathering service and is based on software agent technology. In particular, the goals of the project are:

1. to filter and control the potentially unlimited flux of information from sources to end-users;
2. making information available to people in the appropriate *form*, *amount*, and level of *detail* at the *right time*;
3. to reduce the time spent by the users in knowing regarding: info availability (what, when, where), info structure, info organization, info retrieval services, info access languages and modalities.

The EUROgatherer system will be able to provide the following functionalities:

1. to acquire and retain an interest profile of the user and act upon one or more goals based on that profile;
2. to act, autonomously, pursuing the goals posed by the user irrespective of whether the user is connected to the system where the agent is based;
3. to access a variety of information sources;
4. to create meaningful abstractions of the retrieved documents and classify them appropriately on the basis of their structure and content according to an internal classification scheme, based on user profiles; and
5. to support a relevance feedback mechanism which permits the user to provide the system with feedback on how relevant the retrieved documents are.

The system will collect documents in the following domains in parallel:

- monitoring of frequently changing information sources. The system will monitor at regular intervals URLs that are updated in fixed (or random) intervals for changes. If such changes do exist and are significant then the user will be notified.
- continuous flow of information environment. The system will periodically monitor URLs which generate a continuous flow of information. It will analyze the retrieved documents and select only the proper ones.
- web documents. The system will not search the Web itself, but will utilize existing indexing engines and perform a meta-search in order to discover documents that are, broadly, of interest to the user. Then, the system will further analyze the retrieved documents in order to select those closer to the users preferences.

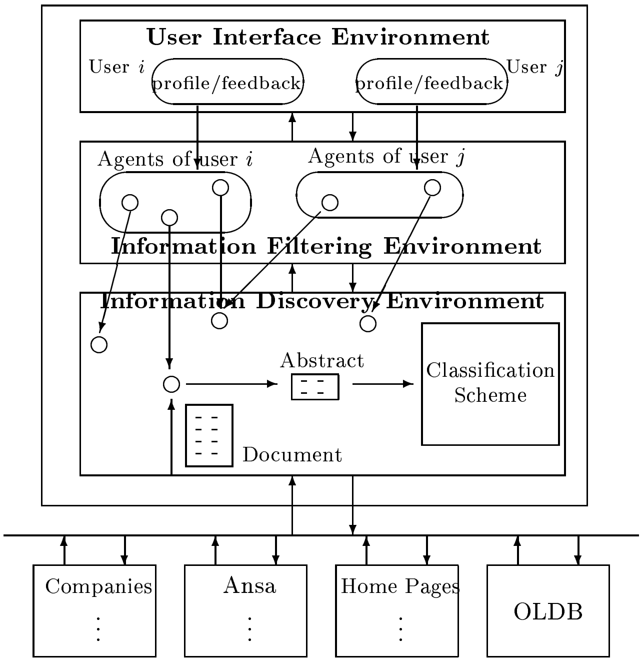


Figure 1: The EUROgatherer system architecture

- on-line Data discovery. The system will access on-line data bases in order to discover data/documents that are of interest to the user.

From the architectural point of view, the project aims at developing an agent-based multilayer system architecture. The system architecture is composed of three layers: the *User Interface Environment*, the *Information Filtering Environment* and the *Information Discovery Environment* (see Figure 1).

Two different species of software agents will be developed: *information filtering agents* and *information discovery agents*. The information filtering agents will be responsible for the personalization of the system and for keeping track of (and adapting to) the interests of the user. The information discovery agents will be responsible for finding, fetching, abstracting and classifying the actual information that the user is interested in. They are utilizing existing Web search engines to find documents (a type of meta-search).

The interactions between the user, the information filtering agents and the information discovery agents are described in terms of a penalty/reward strategy, according to whether the retrieved documents are relevant to the user's needs.

One important aspect of the system architecture is the separation of information filtering and information discovery environments. In the proposed system architecture the personalization of the information, i.e., the information filtering, should be decentralized at the user level, while the information discovery should run on an on-line server. This design choice has a number of advantages:

1. in a multiple user environment, each user will have his/her own set of filtering agents, but they will be able to share their discovery agents;

2. it provides the ability to support real off-line operations; and

3. the introduction of several processing levels between the actual information and the user achieves a greater flexibility in utilizing other novel forms of filtering or other forms of discovery.

Finally, the user interface environment will support the following functionalities:

1. the user profile acquisition by the system;
2. an interactive presentation of the documents retrieved by the system to the user; and
3. the communication of user feedback to the system on how relevant the retrieved documents are.

2 Related literature

1. The Yahoo index. http:// www.yahoo.com
2. The world-wide-web worm. http:// www.cs.colorado.edu/mcbryaan/wwww.html.
3. The Lycos, the catalog of the internet. http://www.lycos.com
4. The metacrawler multi-threaded web search service. http://www.metacrawler.com.
5. K. Decker, V. Lesser, et al., Macron: An Architecture for multi-agent cooperative information gathering. In CIKM Conference, Workshop on Intelligent Information Agents, 1995.
6. Y. Labrou and T. Finn, A semantics approach for kqml - a general purpose communication language for software agents. In Proc. of Conference on Information and Knowledge Management 1994. MIT press, 1994.
7. B. Grosof and al., Reusable architecture for embedding rule-based intelligence. In CIKM Conference, Workshop on Intelligent Information Agents, 1995.
8. A. ORiordan and C. Buckley, An intelligent agent for high-precision information filtering. In CIKM Conference, Workshop on Intelligent Information Agents, 1995.
9. R. Armstrong et al., Webwatcher: A learning apprentice for the world-wide web. In Proc. of the Symposium on Information Gathering from Heterogeneous, Distributed Environments. AAAI Press, 1995.
10. H. Lieberman Letizia, an agent that assists web browsing. In Proc. of the IJCAI-95. AAAI Press, 1995.
11. M. Balabanovic and Y. Shoham, Learning information retrieval agents: Experiments with automated web browsing. In AAAI Technical Report SS-95-08, Proc. of the 1995 AAAI Spring Symposium Series, 1995.
12. B. Sheth and P. Maes, Evolving agents for personalized information filtering. In Proc. of the ninth Conference on Artificial Intelligence for Applications, 1993. IEEE Computer Society Press, 1993
13. N. Belkin and B. Croft, Information filtering and information retrieval. Communications of the ACM, 35, No. 12, 1992.
14. G.S. Jung and V.N. Gudivada, Autonomous tools for information discovery in the world-wide web. Technical Report CS-95-01, School of Electrical Engineering and Computer Science, Ohio University, Athens, OH, 1995.

The Profile Editor: Designing a direct manipulative tool for assembling profiles

Patrick Baudisch
Institute for Integrated Information and Publication Systems IPSI
German National Research Center for Information Technology GMD
64293 Darmstadt, Germany
+49-6151-869-854
baudisch@gmd.de

ABSTRACT

Information filtering systems retrieve documents from document streams according to their users' long-term information interests represented by so-called profiles. The *Profile Editor* proposed in this article allows the interactive, direct manipulative construction of profiles. It takes a set of ranked queries and compiles them into a single profile by cropping and re-ranking the queries' results. The approach of manual profile generation is expected to lead to two advantages: a) Profile generation is expected to be much faster than feedback-based automatic profile generation and b) users' confidence in their profiles should be higher because they are in control of their profiles. The *Profile Editor* is currently being implemented in the context of an Internet TV program guide, in which it will be evaluated during the next months.

Keywords

information filtering, profile, histograms, sliders, direct manipulation, user interfaces, Java

INTRODUCTION

The goal of information filtering systems is to keep users from being flooded with information. Filtering systems remove all items from an incoming information stream that are judged to be non-relevant to users – only those items in the stream that correspond to the long term informational need described in the users' so-called profiles are passed through. See [5] for a comparison between information filtering (or selective dissemination of information, SDI [10]) and information retrieval. Among others information filtering systems have been applied to personal mail and Usenet news [7,8], web sites [2, 13], internet advertising [4].

Profile creation is (not only) an iterative process

Figure 1 shows the model of information filtering as proposed by Belkin and Croft [5]. In this model there are three paths that lead to the *Profiles* node: Creation (top right), outer refinement cycle (thick and dotted boxes) and inner refinement cycle (thick boxes only).

The best explored path of the three is the inner refinement cycle that leads to incremental changes of the profile. The cycle contains three actions ▭ and two documents ▱. The three actions in the refinement cycle are a) *Comparison or Filtering:* Items from the incoming stream are compared with the *profile*. Non-matching items are removed. The remaining items (*retrieved documents*) are presented to the user. These items may or may not contain additional rating/ranking information. b) In the second step (*use and/or evaluation*) the profile system gathers user feedback. In automatic profile generation systems (see below) users are allowed to correct the relevance and/or rating of each item suggested by the filtering system. c) In the third step (*modification*) the *profile* is modified automatically according to the received feedback. The mechanism represented by the inner cycle is also referred to as 'relevance feedback'.

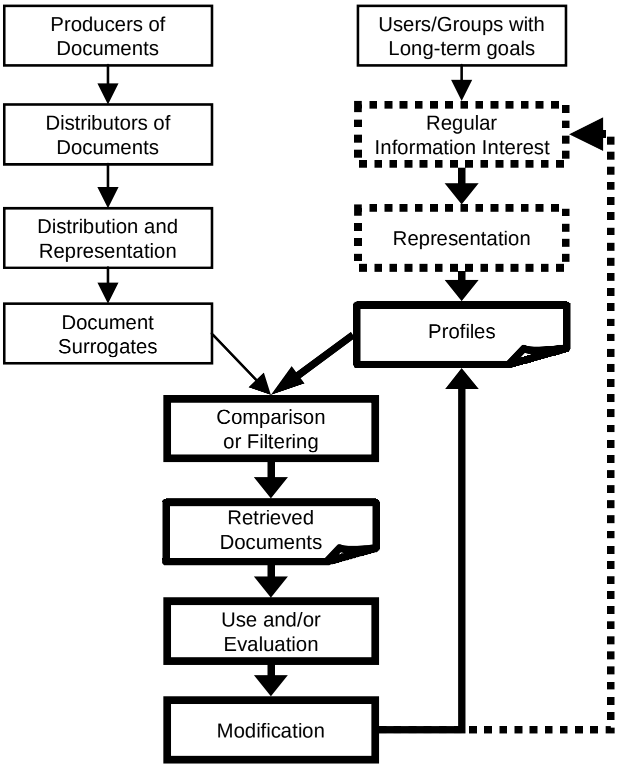


Figure 1: A general model of information filtering according to Belkin and Croft [5]. The boxes with dotted outlines in the upper right describe the first creation of a profile. The boxes with thick frames below describe the inner refinement cycle.

Systems like for example the news filters Gnus [7] and GroupLens [8] implement such an inner refinement cycle.

During the evaluation phase users give feedback about presented items. The systems use this feedback to modify profiles automatically. Users only deal with documents; the profiles never become directly apparent. As a consequence users do not know about the contents of their profiles – they might not even be aware of their existence. Thus there is no profile creation that could represent a regular information interest as stated in Figure 1. And there is no outer refinement cycle that could allow users to communicate how they changed their information interests.

On the one hand the approach of hiding profiles has the advantage of being easy to use. Since the profiles themselves never become apparent, users are not bothered with additional user interfaces or the profile’s internal representation. On the other hand these feedback based profile builders suffer from two limitations. The first limitation is speed. When a profile is created it is either initialized to some stereotype picked by the user or it is even completely empty. In the latter case *all* profile content has to be gathered during the inner refinement cycle. This process takes a lot of time and does not provide a useful profile for quite a long time. The second limitation is the users’ confidence in the profile. If the internal state of the learned profile is not accessible, users can never be sure about the current learning state. This lack of transparency can limit the users’ confidence which in turn reduces the profiles’ applicability in autonomous tasks. Finally the two goals, learning speed and user confidence, seem to exclude each other: Either the learning rate is low and training takes very long, or the learning rate is high and system reactions might be perceived as misunderstandings.

The *Profile Editor* attempts to overcome these limitations by giving users direct access to their profiles. It allows the direct manipulative *creation and modification* of profiles. Its goal is to reduce the number of necessary refinement cycles and to heighten the users’ confidence in their profiles.

A PROFILE EDITOR DEMO SESSION

Before going into detail, let’s take a look at an application example. The following example session shows a possible interaction sequence from the *TV-Online* system [3], a system that assists users in compiling their personal TV schedules.

Andrea assembles her personal TV schedule. She is interested in sports, especially in basketball, where she does not want to miss a single program. She wants to be up-to-date about current information without spending too much time on it. Finally, for recreation, she wants to include some good action movies.

The first thing she does is to select the four genres Basketball, Information, Sports and Action as her favorite genres. In the TV-Online system this is simply done using toggle buttons associated with each genre as shown in Figure 2[1].

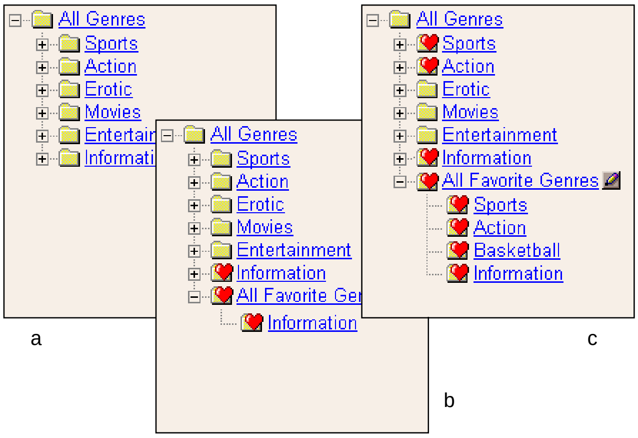
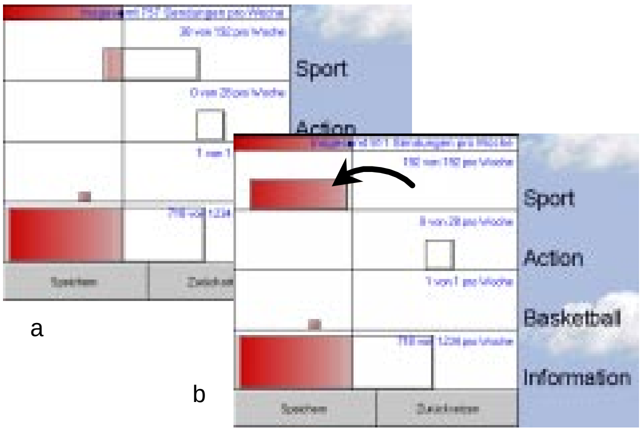


Figure 2: Andrea opens a tree-like menu that contains the hierarchy of all available genres (a) She marks her favorite genre ‘Information’ by toggling the heart icon in front of it. With the selection of the first favorite genre the folder ‘All Favorite Genres’ that holds her new favorite genre appears automatically (b). Finally, she selects the other three favorite genres. The original basketball genre is not visible here – it is hidden inside its parent genre sports. (c)

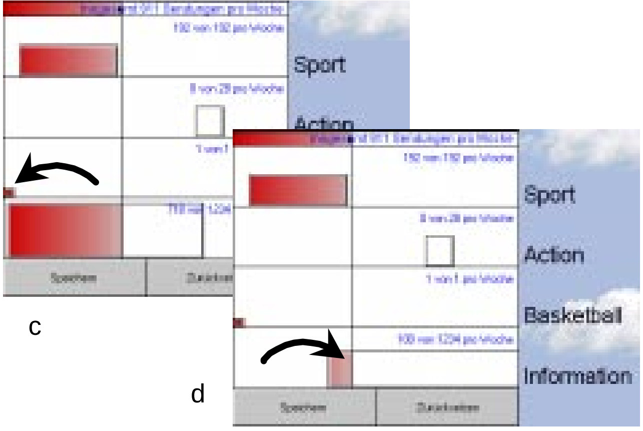
She now has created a personal profile that consists of four genres. She could already query it by selecting ‘All Favorite Genres’ and starting the query process. This would return the union of all programs from the selected genres. See section ‘Initialization’ for details on what Andrea would get and under which ranking. Instead she decides to specify her profile in more detail using the Profile Editor. She invokes it by clicking on the edit button ✎.



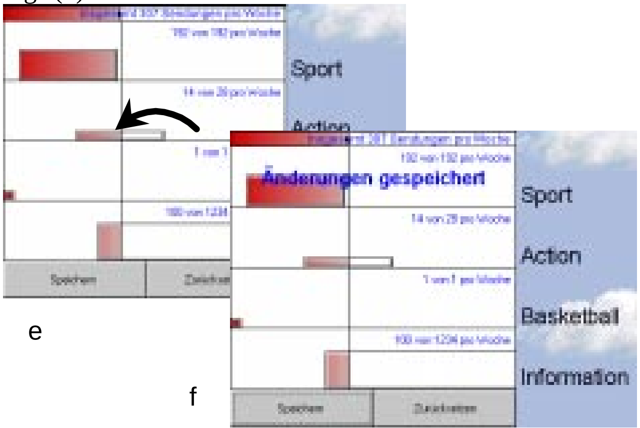
When the Profile Editor is loaded it displays Andrea’s four

[1] In this example the selection of input queries is the creation of the profile, which is the first path to the profile in Figure 1. This mechanism is not understood as being part of the Profile Editor. Different filtering systems might employ different mechanisms of input queries construction or selection. In a filtering system based on a Web search engine the process of choosing queries might be to ‘bookmark’ them.

favorite genres[2] (a). To include all sports programs in her personal schedule Andrea moves the corresponding box completely to the left of the vertical line (cropping boundary) (b).



Then she makes sure not to miss any basketball events by dragging the corresponding box to the utter left. All basketball programs will now be output with a maximum rating (c). Next she reduces the number of selected information programs by cropping them at the vertical line. The remaining hundred programs per week will only get low ratings (d).



Now she moves the better half[3] of the action movies into her selection. She stretches the box horizontally to assign higher ratings to the best action movies. (e) Finally she saves the changes (f). As she can tell from the small text in the containers she now has selected an overall number of 307 broadcasts per week (out of approximately 10,000 on German cable TV).

As she now queries her new profile to get her personal schedule for the current week, the broadcasts returned by her favorite genres are output ranked in the order: All basketball broadcasts, then the top half of all sports programs, then all other sports programs mixed with the better action movies and the top information programs (Figure 3).[3]

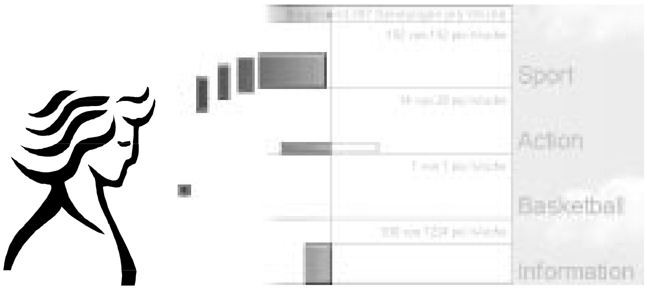


Figure 3: When Andrea queries her profile all items left of the cropping boundary are output ordered from left to right.

BASIC ELEMENTS: HISTOGRAMS AND SLIDERS

Before exactly defining the Profile Editor we will take a look at the basic techniques used. We will start by taking a closer look at sliders and histograms to find out that the draggable boxes demonstrated in the example above represent abstract histograms of query results.

Figure 4 shows a dialog used in a commercial image processor. The dialog allows the conversion of gray scale images into black and white images. The conversion method is very simple in that all brighter pixels are turned to white and all darker pixels are turned to black. The dialog contains a slider that allows the definition of a so-called threshold value, i.e. the luminance value of the darkest color that is converted to white. To assist users in finding an appropriate threshold value the slider is accompanied by a histogram that represents the luminance distribution of an image. Good threshold values might for example be found at local minima around the median of the histogram.

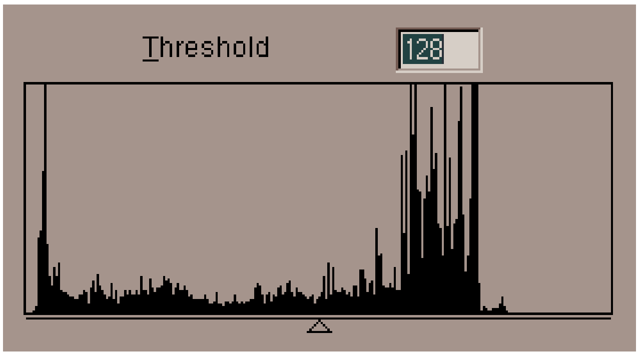


Figure 4: The “threshold” dialog in Adobe Photoshop [1]. The histogram represents the luminance distribution. The little triangle at the bottom is a slider that can be dragged by the user to select a luminance value. The histogram helps in finding useful values to be selection using the slider.

[2] Actually the initial state of the Profile Editor would already be much more appropriate. The shown state was chosen to show all possible interactions. It is a kind of ‘worst case’ initialization. See section ‘initialization’ for the actual initialization.

[3] In the TV-Online example ratings are generated on the basis of other viewers schedules (collaborative filtering [13])

Application to information retrieval

The relation between grayscale and black and white images complements the relation between *rating* and *relevance* in information retrieval. Assuming that items are rated perfectly then there is a boundary that determines which items are still relevant and which ones are not. Like in image processing a histogram slider can be used to select this boundary or threshold. Figure 5 suggests such a user interface component for the information retrieval system Inquery [6] and its graphical user interface Xinquery. Both the original and the proposed widgets visualize ratings of documents sorted by rank. While the bar chart in the original interface represents only twelve documents, the suggested histogram represents about two thousand on the same display area (assuming that each white pixel represents one document). The triangular cursor under the histogram marks the currently selected document and displays its rating and rank. Being able to display the whole range at once provides a quick overview about the amount of returned documents and their rating distribution. Notice the different lengths of the two scroll bars.

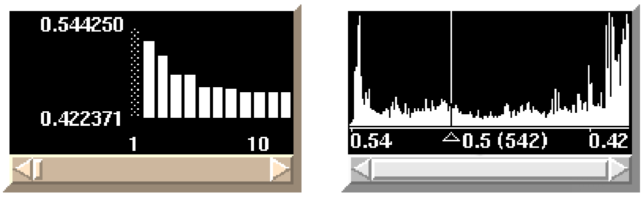


Figure 5: The original Xinquery rating bar chart (left) compared to a widget using a combination of histogram and slider (right). The histogram can represents far more documents than the bar chart. See [15] for more interesting discussion on the Xinquery user interface

To emphasize the relevance property histograms can be colored gradually according to the ratings represented by the individual horizontal positions. The leftmost parts that represent high ratings could for example be rendered red, symbolizing ‘hot’. Parts with only average ratings directly left of the threshold document could be rendered in a pale rose. Parts right of the threshold could be filled with background color to underline that they are not selected.

Application to Information filtering

Applying the combination of histograms and sliders to information *filtering* leads to a number of conceptual changes. In information *retrieval* different informational needs can be processed sequentially. Each informational need is represented by a query which is modified and repeated until the right documents are found (stepwise refinement). When one informational need is satisfied the next one is processed. This approach is not feasible in information *filtering*. Here the data base to search is supposed to be dynamic. Informational needs are expected to be long term interests that exist at least for several sessions. It becomes necessary to hold and maintain several queries at the same time in a so-called profile. Figure 6 illustrates the inclusion hierarchy of profiles and queries.

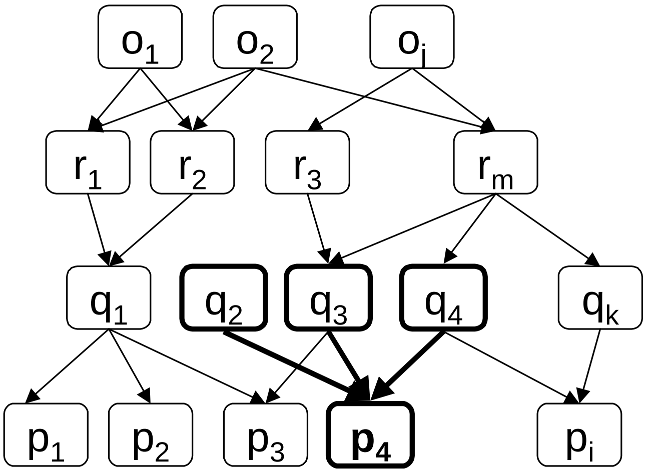


Figure 6: Inference network for information filtering according to Belkin and Croft [5]. O_j are the nodes associated with incoming objects, r_m’s are concept nodes, q_k’s are query nodes and p_i’s represent the profiles. Profiles are collections of queries, The profile p_4 for example includes q_2, q_3 and q_4.

Since a profile consists of several queries, an adapted interface has to contain several histogram sliders, one for every query (Figure 7a). To integrate the results of all these queries into a single output, the ratings of the documents returned by the individual queries have to be mapped to a common domain. To visualize that in the interface, all histograms are inserted into a container that represents this common rating domain (Figure 7b).

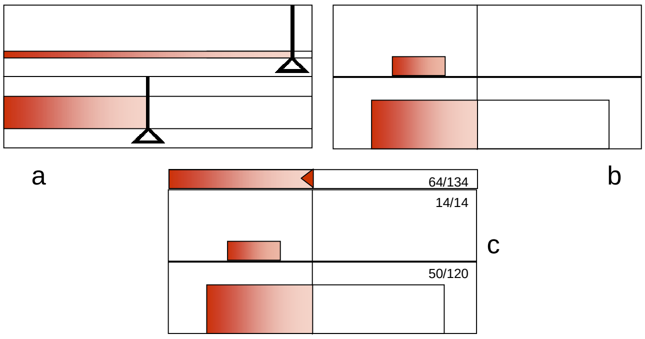


Figure 7: Application of the histogram sliders to information filtering. A profile consists of several queries, each one represented by a histogram slider (a). To integrate the ratings of the different queries into a common space sliders are replaced by a single vertical line called *cropping boundary.* Histograms are moved now instead of sliders (b). The version already presented in the demo session has an extra handle for the cropping boundary and textual information about the number of selected items (c).

Output ratings are now represented by horizontal positions of the surrounding container. The set of threshold sliders now becomes a single vertical line that crosses the whole container. Since this line defines which parts of the result sets will be cut off, we call the line *cropping boundary*. Like the slider, the cropping boundary separates histograms in two subsets: The subset of items that will be returned to the user and the one that will be filtered out. The cropping boundary has only one degree of freedom but has to represent the n degrees of freedom represented by the n sliders before. To accomplish that, *histograms* now have to be dragged instead. The cropping boundary usually stays fixed. Allowing it to be moved as well provides an additional degree of freedom that can be used for influencing the cropping of all queries at once. The next step in adapting histogram sliders to information filtering is to abstract the histograms. The rating distributions visualized by histograms change over time – they might never be the same for two individual runs of the Profile Editor. Therefore no specific histogram can represent all future states of the profile. To avoid misleading information in histograms we use abstract shapes instead. Abstract shapes represent any possible state of a histogram, although not precisely[4]. Of course all advantages related to the display of the concrete rating distribution get lost during abstraction. Different levels of abstraction are possible for displaying and manipulating rating histograms. Figure 8 gives examples. The *abstract* histogram type is very useful in the case that the general type of distribution is known and about constant over time. The *ranked* display is a catch-all: It matches all possible distributions if the rating distribution is replaced by a ranking. Since much rating information gets lost during ranking, abstract histograms should be used instead of normalized histograms whenever possible. Finally, distributions of known type can be given any arbitrary shape by transforming ratings using a continual function. Using this approach any distribution can for example be represented by a rectangle, as we used it for most examples in this article.

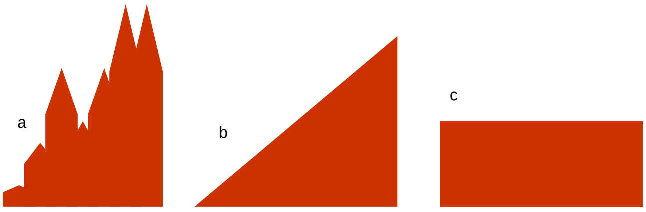


Figure 8: Examples for different abstraction levels of the histogram representation: realistic (a), abstract (b), ranked (c).

Histogram areas have the important function of visualizing the number of presented and selected items. For an example see any figure of the demo session and compare the areas of basketball and information. Based on the area information users are able to estimate how many items they are dealing with and how much effort it is going to take to process the results. Therefore the limited space within the containers makes perfect sense: The overall space left of the cropping boundary represents users' input capacities. By dragging the cropping boundary this space can be customized within the limits of the container[5].

If some queries return very many items while others return only very few the area dynamics may exceed the displayable range. In this case histogram surfaces can be scaled non-proportionally to make sure that even the smallest and the biggest histograms can be easily recognized and manipulated by users. Scaling can for example be done using the following formula:

$$s = s_{\min}\left(\frac{n}{n_{\min}}\right)^{e} \quad \text{with } e = \frac{\log\left(\frac{s_{\max}}{s_{\min}}\right)}{\log\left(\frac{n_{\max}}{n_{\min}}\right)}$$

with s, $s_{\min}$ and $s_{\max}$ being the current, minimum and maximum surfaces respectively, n, $n_{\min}$ and $n_{\max}$ being the current, the minimum and the maximum number of query result items that should be displayed entirely.

PROFILE EDITOR USER INTERACTIONS

The Profile Editor, as implemented in TV-Online, is completely mouse-driven. It supports the following drag and drop interactions:

1. Dragging histograms in the horizontal direction shifts them within the container. Moving histograms to the left increases the ratings of all represented items; moving them to the right decreases ratings. Moving histograms or parts of histograms into the area left of the cropping boundary increases the number of selected items, the opposite decreases the number of selected items. To provide more space for non-selected histogram parts boxes are allowed to stick out to the right.
2. Dragging histograms vertically modifies their aspect ratio. Dragging downwards makes histograms flat and wide, dragging upwards makes them high and narrow. Flat and wide histograms assign a wide spectrum of ratings to the represented items, high and narrow histograms assign similar ratings to represented items. Since histogram deformation can be confusing for novice users, the deforming feature might be omitted in a simplified version. In this case all histograms have fixed aspect ratios. But the additional degree of freedom provided by the change-aspect-ratio feature proved to be quite useful. It allows users to assign arbitrary ratings to the best items while making use of the cropping feature at the same time.
3. Dragging the cropping boundary is a shortcut to modify all queries at once. Moving the cropping boundary to the left decreases the overall number of selected items; moving it to the right decreases it.

[4] To visualize the fact that the histograms in the Profile Editor are not concrete, it might be interesting to give them a less determined shape. Good ideas about so-called non-photo realistic line drawings can be found in [13].

[5] In the TV-Online example there is no such surface restriction: Histograms can be deformed arbitrarily so they can stick out at the top. This is necessary to support the multi-select feature (see section 'Initialization').

The fact that no handles are needed to manipulate histograms makes the interface easy to use. To facilitate the picking and dragging of small histograms any mouse-down event in the whole container (beside those that initiate dragging the cropping boundary) can be used to start a histogram drag interaction.

DEFINITION OF PROFILE AND RATINGS

A profile generated by the *Profile Editor* consists of the position of the cropping boundary b and a set of queries[6] q with their current rating transformation f.

$$profile := ((q, f)^n, b)$$

Each rating transformation maps its query's input ratings r_{in} to global output ratings r_{out}. Output ratings are defined as

$$r_{\text{out}} = f(r_{in}) = r_{in} w + i$$

with w being the horizontal scaling of a histogram and i being the indentation measured from the right. Assuming that both input and output rating are ranged 0 to 1, $w = 1$ assigns the full container width to a histogram. $i = 0$ results in the histogram to be right aligned with its container, $i = 1 - w$ to be left aligned. Inserting this into the profile definition given above this leads to

$$profile := ((q, w, i)^n, b)$$

If an item is returned by more than one input query, the output rating is calculated as the maximum over all r_{out}. Other functions like weighted sums were tested but cannot be discussed here due to space limitations.

The cropping boundary b determines the minimum output rating for items to be returned to the user. The function of the cropping boundary is to remove non-relevant items. The cropping boundary defines the minimum rating for items to be returned to the user. The Boolean variable *output* that determines whether an item is output to the user

$$output := \begin{cases} true & if\ r_{out} > b \\ false & if\ r_{out} \le b \end{cases}$$

with b being the position of the cropping boundary measured from the right of the container. This definition reduces the value range of r_{out} to [b,1]. For many visualizations it will be useful to stretch the output domain to the full range [0,1] by replacing the definition of r_{out} with

$$\mathrm{r}_{\text{out}} = f(r_{in}) = (r_{in} w + i - b) / (1 - b)$$

To support the abstract histogram visualization all queries have to be provided with the average number of returned items and the shape of the typical distribution.

INITIALIZATION

In information retrieval descriptors with low inverted document frequencies are considered more relevant (Law of Zipf, [14, p. 60]). This notion is used to initialize profiles. Queries returning fewer items are expected to deliver more relevant items and are therefore initialized to higher ratings and histograms are placed more to the left.

Additional constraints might be imposed by the application. In the TV-Online system users can create and use profiles without fine tuning, i.e. without using the Profile Editor (see Figure 2), which makes the profile work a kind of multi select[7]. Therefore, all queries have to be initialized as being fully inside the selected range, i.e. left of the cropping boundary. With these initializations users will profit from the indentation created based on the triviality notion even without fine tuning their profile (Figure 9).

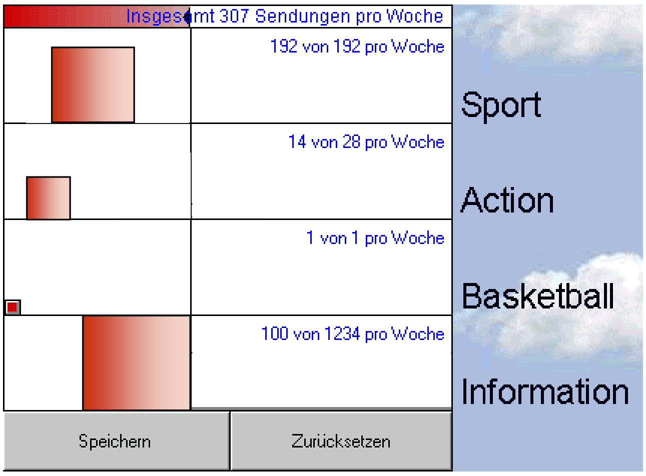


Figure 9: Example of an initialization of newly added queries. Smaller histograms are placed more to the left. In TV-Online all histograms are placed left of the cropping boundary.

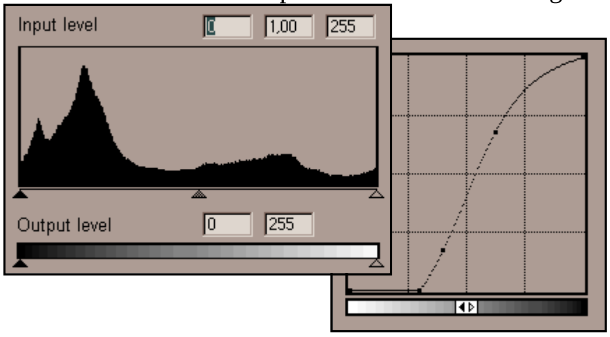
FURTHER RESEARCH

1. Use the Profile Editor on top of Web search engines. The existing service *The Informant* [16] notifies users about newly found pages. The Profile Editor could be used to rank individual queries and to define minimum ratings.
2. Program and test the proposed the histogram user interface component for retrieval systems (Figure 5)
3. Explore and compare different versions of the Profile Editor: Cropping boundary draggable or not, with additional display of number of selected items or not, with extra container for cropping boundary or not.
4. Apply the Profile Editor to image processing. While the Profile Editor maps input ratings to output ratings, gray image filters like the threshold dialog (Figure 4) map input luminance to output luminance. Figure 10

[6] The Profile Editor supports only the definition of the cropping boundary and the transformations — as already mentioned the query set q is expected to be provided by the surrounding system.

[7] As the evaluations showed many users did not want to spend additional work on fine tuning their profiles. In this case it was very important that the profiles worked without the extra effort.

shows two more examples. Can the Profile Editor user interface be used to manipulate multi channel images?



Input level 0 1,00 255

Output level 0 255

Figure 10: Two dialogs from an image processor that map input luminance to output luminance (Adobe Photoshop4.0 [1]).

CONCLUSION

We introduced the concept of direct profile manipulation to fasten profile creation and to increase the users' confidence in their profiles. At the beginning of this article it was presented as an alternative to automatic profile generation as used in systems like Group Lens. But actually the concept of direct profile manipulation is not necessarily opposed to feedback based learning. It seems useful to combine both approaches: Provide a Profile Editor for bigger changes in the outer refinement cycle and to give users more insight into their profile. Use the more convenient feedback learning for incremental changes in the inner refinement cycle. This combination will be the next concept to implement and test.

ACKNOWLEDGMENTS

I thank Dieter Böcker and Ulrich Thiel for their support. Special thanks to Gerrit Voss for the Java implementation of the Profile Editor.

REFERENCES

1. Adobe Photoshop, online at http://www.adobe.com/prodindex/photoshop/main.html
2. Balabanovic, M and Shoham, Y. Fab: Content-Based, Collarorative Recommendation, *Commun. ACM 39,* 6, p. 66–72 (recommend Web sites)
3. Baudisch, P. Designing an Evolving Internet TV Program Guide, *Proceedings of the HCIC '97 workshop,* 19.-23.2.1997, Snow Mountain Ranch, CO. Available Online at http://www-cui.darmstadt.gmd.de/~baudisch/Publications
4. Baudisch, P., Leopold, D. User-configurable advertising profiles applied to Web page banners, To appear in *Proceedings of the first Berlin Economics Workshop,* 24-25 October 1997, Berlin
5. Belkin, Nicholas J., and Croft, W. Bruce Information Filtering and Information Retrieval: Two Sides of the Same Coin? *Commun. ACM 35*, 12
6. Callan, J.P., W.B. Croft, and S.M. Harding: 1992, 'The INQUERY Retrieval System'. In: *Proceedings of the 3*rd *International Conference on Database and Expert System Applications*. Berlin and New York: Springer, pp. 78-83
7. Gnus news reader, online at http://www.aston.ac.uk/lis/as/manuals/xemacs/gnus/
8. Konstan J.A. et al Grouplens: Applying Collaborative Filtering to Usenet News, *Commun. ACM 39,* 6, p77–87, The Grouplens homepage is http://www.cs.umn.edu/Research/GroupLens
9. Netnanny, online at http://www.netnanny.com
10. Parker, K.H. and Soergel, D. The importance of sdi for current awareness in fields with severe scatter of information. *J. Am. Soc. Inf. Sci. 30, 3* (1979), 125–135
11. Pollock, S. A rule-based message filtering system. *ACM Transactions on Office Information Systems* 6, 3 (July 1988), 232–54
12. Robertson, S.E. The probability ranking principle in IR. J. Doc. 33, 4 (Dec 1977), 294–304
13. Schumann, J., Strothotte, Th., Raab, A., Laser, S. (1996), "Assessing the Effect of Non-Photorealistic Images in Computer-Aided Design", *ACM Human Factors in Computing Systems, SIGCHI '96,* Vancouver, April 13-15, 1996, pp. 35-41
14. Salton, G., McGill, M.J., Introduction to Modern Information Retrieval, New York: McGraw-Hill, 1983
15. Shneiderman, B. A User-Interface Framework for Text Searches, D-Lib Magazine, Jan 1997, available online at http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html
16. The Informant, online at http://informant.dartmouth.edu

Using LDAP in a Filtering Service for a Digital Library

João Ferreira(**)
IST – Instituto Superior Técnico
(Universidade Técnica de Lisboa)
Erreur! Source du renvoi introuvable.

José Luis Borbinha(*)
IST – Instituto Superior Técnico
(Universidade Técnica de Lisboa)
INESC – Instituto de Enghenharia de Sistemas e Computatores
Erreur! Source du renvoi introuvable.

José Delgado(*)
IST – Instituto Superior Técnico
(Universidade Técnica de Lisboa)
INESC – Instituto de Enghenharia de Sistemas e Computatores
Erreur! Source du renvoi introuvable.

Abstract

This paper describes how an LDAP directory service can be used to support a filtering service for a digital library. The directory stores and manages profiles of registered users and authors, which are used to implement a filtering service concerned with the submission and change of documents and document annotations, the registration of new users and changes in registered users profiles. The same user profiles are also used to rank results from search tasks as also for user authentication.

1. Introduction

The volume of electronically available information has been increasing in a way impossible to follow by an individual. Filtering services are one possible answer to this problem, and some of those services have been announced, based on user profiles [1]. In ArquiTec, a networked digital library, user profiles are managed in a directory service using LDAP (Lightweight Directory Access Protocol), a standard directory service for Internet. Based on that directory an information filtering service has been built.

The next section briefly resumes the information filtering perspective, presenting a few paradigmatic projects. Section 3 introduces the ArquiTec system, and section 4 introduces the main concepts of the X.500 model, from which LDAP derives. Section 5 describes how an LDAP based solution was used in ArquiTec, and section 6 explains how that directory is being used to implement a filtering system. Finally, the most important open issues are presented in section 7.

2. Information Filtering

Information filtering is an actual subject, with numerous systems appearing and raising important questions. In December 1992 ACM recognized the importance of this new field and published a Communications issue on filtering information. The subject returned again in the March 1997 issue, now focused on a new perspective called «recommender systems». The first filtering systems were targeted for electronic mail and USENET news filtering, but soon those systems were applied to other sources of information, such as the World Wide Web.

Basically, filtering systems use information retrieval techniques in which user queries are replaced by user long term interests, or profiles. These profiles can be created using explicit or implicit methods. Currently, user profiles is one of the richest areas of exploration, specially in the implicit approach (there are experiences, for example, using the time spent reading, analysis of users bookmarks and server log files, etc).

Due to human subjectivity and to achieve better results, several systems involve also humans in the filtering process. For example, in some cases user reactions to the documents are recorded (such as ranking, notes, etc.) and later used to help other users. Those kinds of systems are known as recommender, collaboration or social filter systems.

One of the first historical systems was the TAPESTRY project [2], which coined the term «collaborative system» and raised a new perspective to the problem. TAPESTRY gave two approaches for filtering: *automatic*, where the system evaluates what is interesting to the user, and *social*, where users help each other.

Table 1 summarizes a few paradigmatic systems developed until now or under development, emphasizing the users profile and matching techniques.

Sift and Newsweeder represent two examples of *automatic* filtering systems. The basic difference between them is the way profiles are defined, where Newsweeder uses also an implicit method based on past user experience. *Automatic* filter has had success only in very simple systems. The main problem is that it has to deal with the issue of automatic creation of representatives of documents (or surrogates), a complex task even for well-defined areas.

(**) Work supported by the JNICT fellowship PRAXIS XXI/BD/5968/95

(*) Work partially supported by the JNICT contract PRAXIS/2/2.2/TIT/1667/95

Table 1: Some paradigmatic filtering systems

As examples of *social* filtering systems we have Grouplens and ReferralWeb. Those systems are in general more successful than the automatic ones, but unable to provide information in documents that have never been read. Another weakness is the problem of finding the correct tools to keep out (or to minimize the effect) of disruptive users (such as, for example, users who are not really collaborative but only interested in giving high rates to themselves or related friends).

System	Information Source (IS)	Profile		Matching		Remarks
		Explicit	Implicit	Tecniques	Arguments	
Grouplens (1992)	Usenet	Numeric Vector	Numeric Vector (reading time)	Cosine measure	(user profile) versus (users profiles)	Collaborative filter system
Sift (1994)	Usenet	Keywords lis	-	Boolean	(IS) versus (user profile)	Filter system
Newsweeder (1994)	Usenet	Numeric Vector	Numeric Vector (user history)	Cosine measure	(IS) versus (user profile)	Content based-filter system
Fab (1994)	Web	Numeric Vector	Numeric Vector (user history)	Cosine measure	(IS) and (users profiles) versus (user profile)	Collaborative and content-based filter system
ReferralWeb (1994)	Web	***mention of a person or a document***			***(user profile) versus (community profile)***	Collaborative and social filtering

Concerning the matching techniques, two main approaches have been tested: (i) to match the profile against other profiles and to choose the information in the nearest one (collaborative) or (ii) to match against community standard profile and to use the nearest standard to get information (social). Fab is a system that tries to combine both approaches.

3. ArquiTec

The ArquiTec project aims to develop a digital library for the Portuguese scientific and research community [3]. It started in the beginning of 1997, and a first phase will end with a working prototype, scheduled for public release in the first quarter of 1998.

ArquiTec is accessible over the Internet, through a WWW interface. It provides access to different kinds of technical documents (such as papers, reports, theses, dissertations, etc.), in different fields of knowledge, while special services will also be provided to the community.

The system was conceived around three main entities, as shown in Figure 1: *documents, users* and *concepts*. Informal and formal documents exist in local repositories, managed by a structure of distributed servers based in the NCSTRL technology [4]. To address the problem of long term preservation, the Portuguese National Library will maintain a PURL service [5] and a central official archive with a copy of selected formal documents (such as thesis and dissertations).

ArquiTec users can be authors, readers, or both. Users are managed in a global X.500 like directory [6], where their identity, contacts, affiliations and a special profile are registered. Anonymous access is possible for search, browse or even retrieval, but users are always suggested to identify themselves for profile management.

The concept space, or ontology, is based in the integration of possible multiple statistical and formal thesauri, as well as user contributions. Two important components of this space are user and collection statistical thesauri, created from the document collections and also from the user directory (profiles). In that sense our thesauri perform functions well beyond their usual roles as auxiliary tools for classification and search. Matching these thesauri with the collection makes it possible to identify document clusters, for example, but it makes also possible to identify virtual user communities (defined as groups of users sharing common interests).

Documents, users and concepts are interactive and dynamic entities, which means that they can change over time. For example, documents can have new releases or attachments (submitted as annotations), users can become interested in new subjects, new subjects can be included in concept space, new relationships can be established between existing subjects, etc. Indexes, user profiles and the relations between documents and users (authors or just readers) associate these entities among themselves.

Users are identified in ArquiTec by their interests and contributions, which relate to subjects in the concept space (likewise for documents). In ArquiTec users are viewed not only as authors and patrons but also as important sources of information, with their profiles becoming part of the contents. Profiles serve also to provide special services to the users, such as filtering (automated notifications) and ranking of search results.

Conceptually, a *catalog* makes it possible to explore, in an integrated perspective, the above six concepts (comprising the three main entities and the relationships between them). In that sense it becomes possible and has an equivalent meaning, for example, to search for documents or for users related to a specific subject (in an integrated perspective, it is also possible to search for both users and documents related to specific subjects, and so conceptually «sharing common interests»).

4. X.500 and LDAP

ArquiTec uses an X.500 directory in an LDAP implementation.

X.500 is an OSI directory service, which defines an information model, a namespace, a functional model and also an authentication framework. An X.500 directory is based on entries, which are collections of attributes as defined in RFC 1779 [7]. Each entry has a type (or class), typically defined by one or more mnemonic strings, and can have one or more values.

The attributes required and allowed in an entry are controlled by a special object class attribute in every entry. The information is supposed to be structured in a tree, accessible by servers possibly distributed over a network.

As shown in Figure 2, at a top level there are entries representing countries, below that there are entries representing national organisations, and so on. At the lowest level it is supposed to find entries representing any desired class of objects, such as people, computers, printers, etc.

X.500 defines the Directory Access Protocol (DAP) to access the service, a full, complex and heavy OSI protocol supporting operations in three areas: search/read, modify and authenticate. The search is possible at any level, based in a filter query involving attributes and returning requested attributes from each matching query.

The problem of the excessive complexity of the DAP protocol has been addressed by the Network Working Group of IETF, which has been proposing the Lightweight Directory Access Protocol (LDAP) as an alternative for the Internet.

LDAP is a client-server protocol that runs directly over TCP/IP, and it was conceived to remove some of the burden of X.500 access from directory clients, such as taking out some of the less-often-used service controls and security features.

LDAP is being positioned as the directory standard for the Internet, with leading industry players like Microsoft, Netscape, IBM, Lotus, Novell and Banyan supporting it or intending to support it in the near future [8]. There are also plans to develop LDAP access for several database and index machines, such as Glimpse, for example).

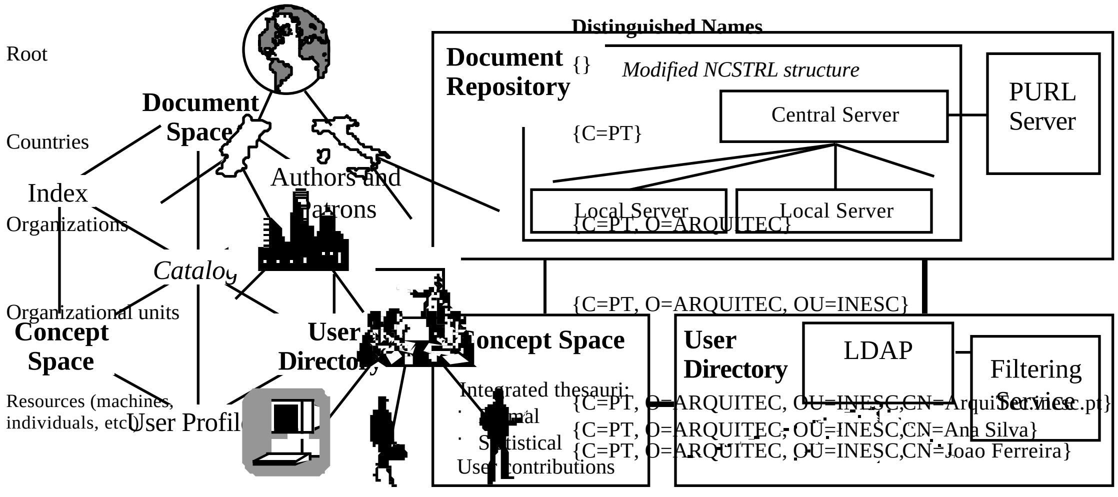


Figure 1: Main entities and block architecture of ArquiTec

Figure 2: Typical X.500 structure

Generic attributes	Profile attributes
Name	**Explicit Fields:**
Institution	List of interesting subjects
User identifier (ArquiTec)	List of non interesting subjects
Password	**Implicit Fields:**
Password tip	List of identifiers of archived documents which the user has authored
Email address	List of subjects of documents which the user has authored
Telephone	List of identifiers of submitted annotations
Fax	List of identifiers of retrieved documents
WWW home page	List of subjects of retrieved documents

Table 2: ArquiTec user entry in the user directory.

5. LDAP in ArquiTec

ArquiTec uses the Directory Server package, an LDAP implementation from Netscape and based in an original work from the University of Michigan [9]. This LDAP implementation has three main components:

Server: our server runs on a Unix machine as a stand-alone daemon.

Client library: a powerful C language API for accessing and using LDAP, with LDAP clients and a backend to handle database operations [10].

Gateway: a special WWW interface is available for directory and server administration.

Users access ArquiTec in one of two modes: anonymous or identified. Identified users have profiles composed of explicitly provided data (their explicit interests) and data implicitly extracted from the history of their interactions with the system (such as submitted and retrieved documents).

ArquiTec users are managed in a structure such as presented in Figure 2, where each user entry has a list of fields as presented in Table 2.

At the moment, the user directory is implemented in only one server. However, to provide flexibility and fault tolerance it will be distributed and replicated it in the near future by other servers within the national academic network (a feature supported by LDAP).

1. Filtering in ArquiTec

In ArquiTec the filtering service follows both the *automatic* and *social* approaches. It is a social system because document classification gets richer with annotations submitted by users. It is also an automatic system because it automatically matches new documents and annotations with the existing user profiles and new profiles with the existing documents.

More generically, user profiles serve three main purposes in ArquiTec, as shown in Table 3:

Filtering: profiles are used to provide an information filtering service, supported by electronic mail, through which users can receive automatic notification of new events.

Searching: profiles can be used to rank search results, for example to highlight documents that best match user's interests (but ranking will never hide or restrict the access to other documents that also match the queries).

Retrieval: the access to different kinds of documents or to special user information can depend of the user profile. This is a scenario not yet implemented in ArquiTec, where privacy protection concerns have to be taken in account, since it requires defining profiles fields not controlled by the user but by an administrative authority (in the current scenario user profiles are public and fully controlled by the users).

The filtering service tracks five kinds of events:

Notification of new documents: any user whose profile matches the classification of a new document is informed about it (to submit a document, a metadata form has to be filled).

Notification of changes in stored documents: if a new version of a document is submitted, users that, for example, had retrieved that document, will receive a notification.

Notification of new annotations: any user whose profile matches a new annotation will be notified about it (in

Event sources	User profiles usage in ArquiTec		
	Filtering Service	Information Search	Information Retrieval
Documents	- New documents - Document changes	Ranking of query results	Control Access
Annotations	- New annotations		
User Profiles	- New users - Changes in profiles		

Table 3: Usage of user profiles in ArquiTec

fact, an annotation in ArquiTec is just a document metadata form, similar to the form filled in the submission of the document).

Notification of new users: when a new user is registered, users with similar profiles will be notified.

Notification of changes in user profiles: when a user profile changes, users matching the new profile will be notified.

User profiles can be used also to rank search results, giving more relevance to results that best match the profile of the user. For this task, it is also possible for the user to choose to identify him/herself with a virtual profile created by the system, instead of its own.

From the user directory it is possible to identify groups of users with similar profiles, and so to create virtual profiles of possible communities. In the future, this feature will be exploited for collaborative services, such as mailing lists (automatically created).

1. Future Work

ArquiTec is work in progress. The structure of the user profiles still need to be tested and tuned (it was defined until now in a mixture of implicit and explicit methods). Access restrictions to information (documents and user profiles) will be also implemented based in different criteria, namely in administrative fields in the user profile.

Work has to be done yet in the conceptual space based on the collection statistical thesauri and user directory. An important open issue here is the creation and maintenance of authority lists, vital to control the integration of thesauri. Finally, an exciting issue is the development of strategies for the (semi-)automatic identification of user communities and the conception of new services based on that perspective.

References

[1] Resnik, P.; Varian, H.R. (1997). **Recommended systems**. Communication of ACM, March 1997, Vol. 40, N. 3

[2] Goldberg, D.; Nichols, D.; Oki, B.M.; Terry D. (1992). **Using collaborative filtering to weave an information TAPESTRY** Communication of ACM, December 1992, Vol. 35, N. 12.

[3] Borbinha, J.L.; Ferreira, J.; Jorge, J; Delgado, J. (1997). **A Digital Library for a Virtual Organization.** Proceedings of the **Erreur! Source du renvoi introuvable.**.

[4] Davis, J.R. (1995). **Crating a Networked Computer Science Technical Report Library**. D-Lib Maganize, September 1995. *Available on-line in 27 September 1997 at http://www.dlib.org/september95/ 09davis.html*

[5] Weibel, S.; Jul, E. (1995). **PURLs to improve access to Internet**. OCLC Newsletter, November/December 1995, 19. *Updated version available on-line in 27 September 1997 at http://purl.oclc.org/ OCLC/PURL/SUMMARY*

[6] CCITT (1988). **X.500 The Directory: Overview of Concepts, Models and Service**. CCITT Recommendation X.500, 1988.

[7] Yeong, W.; Howes, T.; Kille, S. (1995). **RFC 1777: Lightweight Directory Access Protocol**. IETF Network Working Group, March 1995. *Available on-line in 27 September 1997 at http://ds.internic.net/rfc/rfc1777.txt*

[8] Cooper, J.; Ratcliffe, N (1997). **The role of LDAP and X.500**. Data Connection, August 1996. *Available on-line in 27 September 1997 at* **Erreur! Source du renvoi introuvable.**

[9] Howes,T.; Smith, M. **LDAP Programming Directory-enabled Applications with Lightweight Directory Access Protocol.** Macmillan Technology Series (1997)

[10] Howes, T; Smith, M. (1995). **RFC1823: The LDAP Application Program Interface**. IETF Network Working Group, August 1995. *Available on-line in 27 September 1997 at http://ds.internic.net/rfc/rfc1823.txt*

SOaP: Social Filtering through Social Agents

Hui Guo, Thomas Kreifelts, Angi Voss

GMD-FIT.CSCW
German National Research Center for Information Technology
Institute for Applied Information Technology
Schloß Birlinghoven
D-53754 Sankt Augustin, Germany
Email: {firstname.lastname}@gmd.de

Abstract

The Web is becoming the premium source of information for a growing number of people. As a result, information overload arises as a problem of extracting useful information. Information gathering on the Web has become a time-consuming work. As an emerging technique for dealing with this problem, collaborative filtering (also known as social filtering) can improve retrieval precision and reduce the amount of time spent. In this paper we propose a social filtering system consisting of various types of agents which mediate between different people, groups and the Web. Agents work on behalf of their clients–users or other agents–under the specified security and/or privacy constraints. They interact with each other and allow people to cluster the URLs, rate and annotate the Web pages, and share the recommendations. Agents could also find people and groups with similar interests, bring people together to form groups and allow them to work within various groups to exploit the collected bookmarks. Eventually, the system could contribute to the social construction of knowledge on the Web.

Introduction

Less than five years old, the concept of collaborative filtering [Shardanand & Maes'95] has already spawned dozens of research prototypes, experimental proprietary systems, and even a few commercially available systems. Our project at the CSCW group of GMD was set up to create an open distributed platform that supports various Web-situated social applications through a wealth of interacting software agents. Ultimately, our agents should support the social construction and evolution of knowledge by communities of people wired on the Web.

By an "agent", we mean an autonomous software process which acts on behalf of a client. We speak of "social agents" when the agents support the social relationship between their clients. Social agents can play several roles:

- As workhorses, they can use the idle time to do repetitive routine work, and thus reduce their clients' overhead and increase the cost/benefit ratio of their service.
- As representatives, they can learn their clients' preferences and changing interests, present them to others while protecting the privacy, thus making personalization feasible.
- As mediators, they can exchange information, match interests, negotiate on behalf of their clients, and bring people with similar interests together so as to facilitate social affiliation.

The first application SOaP built upon the platform is to provide people with "live bookmarks" by combining content-filtering search engines with collaborative filtering techniques from recommender systems [Resnick & Varian'97] to exploit people's assessments of Web pages. It mines the Web for bookmark pages in order to acquire a critical, initial mass for recommendations. To account for motivational factors, it provides groups as contexts where people may feel more compelled to exchange and discuss assessments.

Related to our work are groupware systems that help people to share bookmarks and annotations, Web-based software agents [O'Leary'97], and recommender systems [Resnick'97]. These systems deal with a wide variety of information from the Internet, e.g. URLs, hyperlinks, annotations, queries, bookmark folders, emails, or newsgroup articles. They largely differ in the way that information can be shared. In some annotation sharing systems (JASPER [Davies et al.'95], users can or have to indicate explicitly how to distribute their information. In Answer Garden 2 [Ackerman & McDonald '96], users can specify how to escalate their questions. Thus, the developers claim, the problem of missing context can be reduced because people with a common background can be consulted first. Also, they argue that agents as mediators between persons can side-step certain social barriers. In some others systems, information can be distributed based on implicit criteria: content of information or ratings from human users. Letizia, WebWatcher, JASPER and FAB allow users to specify their interests in terms of keywords, and propose information which matches this description.

The following systems minimize the user's overhead. They extract URLs from sources that exist independently on the Web. PHOAKS [Terveen et al.'97] extracts citations of URLs from newsgroup articles, categorizes them, and recommends more recently or more frequently mentioned ones. While PHOAKS' output is not tuned to a particular user, GAB [Wittenburg et al.'95] and Siteseer [Rucker&Polanco'97] take the user's personal bookmark folders as an implicit declaration of interest. Both compute overlaps with other people's bookmark collections. GAB takes bookmark folders from explicitly selected users, while Siteseer takes bookmarks from a large collection of anonymous users. In contrast to our system they do not interpret the content of the folder titles. Yenta [Foner'97] is planned to consist of decentralized user agents, only. They will analyze the user's outgoing email and try to form clusters. User agents will compare each others' clusters and recommend users with a good overlap. The agents also exchange their acquaintances to find new candidates.

Compared to the systems above, SOaP is unique in its combination of functionality, and it can be easily extended to offer more, such as recommending users with similar interests. Within this system, users have personal workspaces in which they can input queries and get bookmarks back, and are allowed to cluster, annotate and rate these periodically updated bookmarks, as well as form groups to share bookmarks and annotations. The system can match users' interests against each other in terms of different levels of context (query, task, and group), make recommendations accordingly, and make it possible to aggregate knowledge out of user ratings and annotations. The system functionality is basically supported by several types of communicative agents: user agent, task agent, query agent, recommender agent and search agent. These agents have their own knowledge about the user or the environment, and interact with each other through a set of uniform performatives to exchange information in order to achieve their specific tasks. With SOaP, we hope to tackle two important problems, missing context and cold start. Missing context can be supplied by relating URLs to topics and groups. The cold start problems is avoided by starting with published bookmark collections, by using search engines to introduce new material without any personal overhead. In addition, the scalability of SOaP is expected to be guaranteed by distributing required functionality over individual agents which are distributed over networks.

In the following, we give an outline of our system comparing it with others with respect to well-known issues such as incentive and privacy, group context, etc. We also briefly introduce the underlying infrastructure and system status.

Agent-based social filtering for recommendations

In our system, information is filtered by communicative social agents which collect human users' assessments and match users' interests to derive recommendations. With agents, it is possible for users to find relevant bookmarks regarding specific topics, find people with similar interests, find groups with similar topics, and also to form groups for direct cooperations.

User interface and agents

Users interact with the system via dynamically created HTML pages. As shown in figure 1, every user has a personal workspace where all the user's tasks, queries, bookmarks and annotations are kept from unauthenticated access. One root page is linked to the pages of those groups the user belongs to, to a summary page of all groups, and to the pages of the user's tasks. A task page contains the queries issued for the task, the results obtained from the queries or bookmarks dropped by the user, along with their ratings and annotations. There are forms to formulate queries, to import bookmarks and to input annotations. A group page lists the members of the group and is linked to the group's task pages. Task pages of groups are essentially shared user task pages. As a difference, they indicate disagreements between ratings, and the annotations serve as a record of the discussion process in the group. The group summary page (not shown in the figure) summarizes the information about all groups along with their members, hotlists, and tasks. In this system, agents serve three purposes as illustrated in figure 1: they construct and maintain pages of the user interface; they wrap or manage databases which are accessible for retrieval; they perform retrieval subtasks.

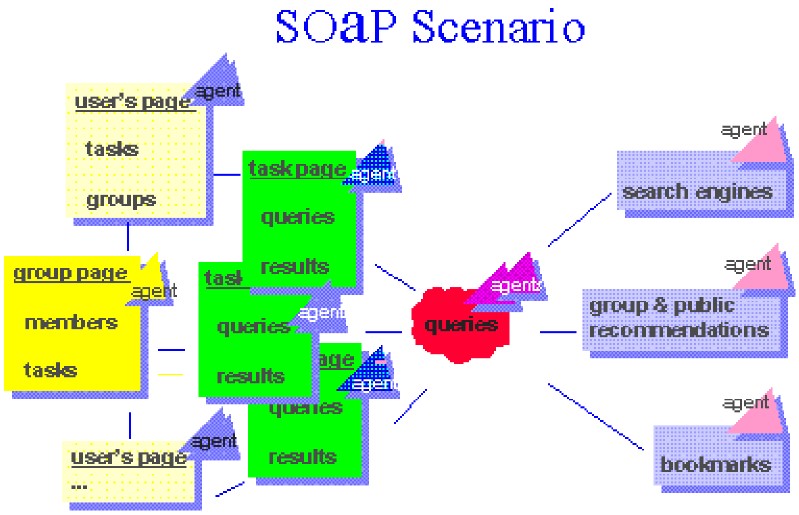


Figure 1: SOaP agents operate between users, groups and the Web.

SOaP Agents

In order to perform the services, agents in our system use knowledge about users, groups of users, the topics that are relevant to a user, the URLs that a user considers relevant to a topic, and a user's assessments of a URL, e.g. his or her ratings and annotations, in the context of a particular group or in connection with a particular topic. According to our design principle, this knowledge should be obtained without effort on the part of the user, or else it should be optional.

Agents in SOaP are specialized with regard to behavior and function. Each type of agent has specialized task to accomplish and plays different roles in the overall system. They interact with each other by exchanging messages of a certain type ("performative"). The interaction between them are conceived of as a conversation. Such conversation patterns may be formalized as finite-state-machines or in a distributed environment as high-level Petri nets following [Kreifelts & v. Martial'90]. The FSM specification of each conversation can be described by conversation tables which specify the state transitions as well as the message to be sent/received in particular states. They may be used to formally verify that the conversations is free of deadlocks even in the presence of message delay and mixed initiative of the conversation partner (not a simple turn-taking protocol) [Woetzel & Kreifelts'89]. The interaction between agents varies in complexity and duration, and principally can be captured by this concept of a conversation, Also it is possible to introduce new ways for agents to interact by specifying new conversation types, e.g. for the interaction between user agents and recommender agents if one wants to migrate certain tasks from task agents into user agents.

Within the system, as shown in figure 1, each registered user has his/her own unique user agent, the user's permanent representative which accepts the user's queries, then initiates task agents which in turn distribute queries to individual query agents. These query agents request services from a search agent—a service wrapping popular search engines like AltaVista, Infoseek—and a recommender agent which implicitly collects all ratings and annotations from users and performs matching and recommending. After receiving recommendations from these services, query agent, task agent and user agent will cooperatively merge, sort, filter, combine, cluster these information by relevance and present tailored results to users. By using this system, users can iteratively input and refine queries, rate the resulting URLs between -2 (very bad) and 2 (very good), make annotations, review recommendations from others who have similar interests, and gradually obtain a valuable bookmark list with respect to a specific task context which groups all queries, bookmarks, ratings, and annotations. This bookmarks list is alive and grows, for the recommender will autonomously push the new recommendations to each registered query agent which presents them to the user accordingly.

Since each user has his/ her own user agent and interacts with this agent for collecting bookmarks relevant to the topics he or she is interested in, this agent can capture the user's changing interests explicitly or implicitly. Also the user's interest is expressed in the form of queries in a specific task context, and the user's relevance judgment of certain bookmarks is made with respect to the current context. The user agent can derive s user profile out of all these information which can dynamically reflect the user's preference and interest.

In our system, a recommender agent served as a common repository for a society of query agents and task agents. It registers each user's tasks and queries, and stores feedback from users in the forms of ratings and annotations. Compared with a search engine, it applies both content-based and collaborative filtering in order to recommend. The content is filtered by comparing the query with each bookmark which is described and indexed by keywords with their weights. For collaborative filtering, only URLs are selected which have obtained a high combined rating from other users. Recommendations are made based on calculating similarity of stored tasks and results. Task agents register their tasks with the recommender agent, which clusters the tasks (by defining neighbours of each task) and makes ratings and annotations automatically available to those neighbouring task agents within certain cluster. The recommender agent can also only recommend other task agents worthwhile negotiating with, which may be necessary for task agents which interact directly with each other in certain application like buying and selling on the Web.

Group context

To provide a richer and more constrained context for collaborative filtering, we propose a group agent, which store queries, results, ratings and annotations of group members like a recommender agent A group agent allows users to annotate and share bookmarks, matches users' interest and recommends highly-rated bookmarks along with annotations. By cooperating with other information resources like search agents, recommender agents, and other public group agents, a group agent behaves like a space for users to share recommendations and construct personalized views of group bookmark collections. Since users can only join a group by invitation, their identity is visible with regard to ratings or annotations within the group. This prevents non-serious anonymous ratings and ensures better recommendations.

The members of a group form a community; their recommendations will help each other so that providing many and good recommendations (evidenced by other users) can improve their prestige in the group. It seems that cold start is less of a problem in suitably selected groups and social motivations can stimulate substantial personal effort with a group. Last but not least, sharing recommendations within a group allows to detect trends faster than through other communication channels. Trends in groups may be interesting not just for group members, and more generally, some information should be allowed to leak out of a group so that other people may get a chance to join the group. Therefore our system not only allows to use a group's ratings for public recommendations, but allows to suppress selected annotations by specifying what kind of information should be *public / private* to non-members.

Related issues

A well-known problem of recommender system is the cold start problem. A recommender system can produce good recommendations only after it has accumulated a large set of ratings. In our system, a bookmark agent is introduced which collects and exploits publicly available bookmark pages by transforming them into the internal recommendation format for use by the recommender agent.

In order to protect the users' privacy appropriately, each information object which may be transmitted to other users as part of a recommendation has privacy-related access control. A typical example is the annotation. A user can make annotations anonymously, which means the user name won't be visible to whoever might get this annotated recommendation. Privacy control is also defined with respect to context, e.g., if a task is made private, all the queries under this task will be private by default as well.

Also in some recommender systems, there exists a "vote early and often" phenomenon, resulting in recommendations based on faked ratings. Such user actions are detected and inhibited by the user agent in our system: all the highly-rated bookmarks are sent to the public recommender agent only once, and repetitive ratings of the same bookmark will be simply ignored by the user agent. The user agent could also fetch each bookmark page and perform content-based filtering to filter out apparently inserious ratings, however this is not our focus.

The incentive (or cost/benefit) problem is addressed in the system in two ways. First, we keep the rating overhead for users low by interpreting users' actions as implicit votings: our system allows user to discard a URL or to annotate it without rating, and interprets this as negative or positive feedback, respectively. Secondly, our system provides a better quality of information retrieval than public search engines by exploiting human judgements and recommendations, so there is a basic benefit with no cost of rating and annotating. For mutual sharing of costs (and benefits), users are encouraged to rated and annotate, especially in the context of specific user groups.

SOaP implementation

As an application that allows users to assess Web pages and recommend them to other users and groups, SOaP is implemented on top of an open infrastructure. This infrastructure is a distributed platform for interacting software agents which provides a runtime environment with basic functionality, such as a unique agent naming/addressing schema, a message-based agent communication mechanism, fault recovery, persistency of agents, multiple agent accounting and a distributed directory service for agents. This platform is composed of the kernel layer-agent engine that hosts multiple agents and represents the basic runtime environment with kernel services like agent creation, termination, messaging, etc.-and a service layer that implements system services like the naming and alarming services. Both application and infrastructure are implemented in Java. The overall layered architecture is shown in figure 2.

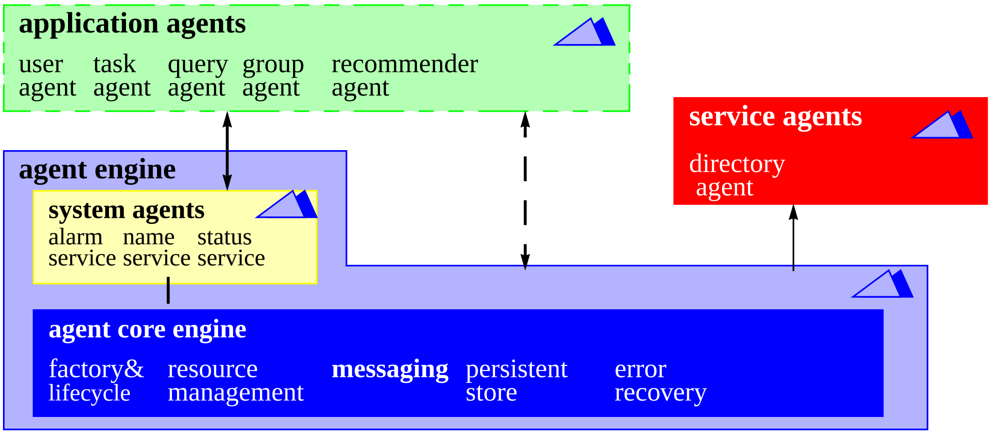


Figure 2: SOaP - the layered agent architecture (one host)

Status and future work

The first prototype released for tests within our research group at GMD in Septemeber'97 included agents which implemented the user interface, tasks and queries, and agents providing information by contacting search engine or collecting recommendations. This prototype is intended for demonstration, exploratory use, and evaluation in cooperation with an industrial partner from the oil business. Prospective users are members of project teams operating in oil field development. Team members are usually spread around the world, and may belong to several teams at the same time. Information retrieval and exchange is central to their work, and it has to occur both in similar areas or using similar techniques.

For the next release, we consider to design agents which provide interface to legacy information system and shared workspace system like BSCA [Bentley et al.'97]. In the context of SOaP, more problems arise such as matching, clustering, use of thesaurus to capture group-specific terminology, and creation of summary queries as operational definitions of tasks, and need further investigation in future research.

References

[Ackerman et al.'96] Ackerman, M.S., McDonald, D. W. "Answer Garden 2: Merging organization with collaborative help," in Proc. CSCW'96 (Cambridge MA, 1996), ACM Press, New York NY, 1996, pp. 97-105.

[Bentley et al.'97] Bentley, R. Appelt, W. Busbach, U., Hinrichs, E., Kerr, D., Sikkel, K., Trevor, J., Woetzel, G. "Basic support for cooperative work on the World Wide Web," Int. J. Human Computer Studies 46('97), 827-846

[Davies et al.'95] Davies, J., Weeks, R., Revett, M. "JAPSER: Communicating information agents for the WWW," in Proc. 4th World Wide Web Conf. (Boston MA, Dec.1995), World Wide Web Journal Vol. 1, 1, O'Reilly, Sebastopol CA, 1995, pp.473-482

[Foner'97] Foner, L. N. "Yenta: A multi-agent, referral-based matchmaking system," in Proc. Agents'97, 1st Int. Conf. on Autonomous Agents (Marina delRey CA, Feb. 1997)
http://lcs.www.media.mit.edu/people/foner/Yenta/overview.html

[Kreifelts & v. Martial’90] Kreifelts, Th., v. Martial, F. "A negotiation framework for autonomous agents," in Y. Demazeau, J.P. Mueller (eds.) Decentralized A.I. Vol 2, Proc. MAMAAW’90 2nd European Workshop on Modelizing Autonomous Agents and Multi-Agent Worlds, North-Holland, Amsterdam, 1990.

[Resnick & Varian’97] Resnick, P., Varian, H.R. “Recommender Systems,” Comm. ACM 40, 3 (1997), 56-58

[Rucker et al.'97] Rucker, J., Polanco, M. J. "Siteseer: Personalized navigation for the Web," Comm. ACM 40, 3(1997), 73-75

[Terveen et al.’97] Terveen, L. Hill, W. Amento, B. McDonald, D. “PHOAKS: A system for sharing Recommen dations,” Comm. ACM 40, 3(1997), 59-65

[Voss & Kreifelts '97] Voss, A., Kreifelts, Th. "SOaP: Social Agents Providing People with Useful Information," in Proc. Int. ACM SIGGROUP Conf. on supporting group work, ACM, New York NY, 1997, pp-291-298.

[Wittenburg, et al.'95] Wittenburg, K., Das, D., Hill, W., Stead, L. "Group asynchronous browsing on the World Wide Web," in Proc. 4th Int. World Wide Web Conf. (Boston MA, Dec. 1995), World Wide Web Journal Vol. 1, 1, O'Reilly, Sebastopol CA, 1995, pp.51-62.

[Woetzel & Kreifelts'89] Woetzel, G., Kreifelts, Th. "Deadlock freeness and consistency in a conversation system," in B.Pernici, A. A. Verrijn-Stuart (eds.) Office Information Systems: The Design Process, Proc. IFIP WG 8.4 Work. Conf. On Office Information Systems: The Design Process, North-Holland, Amsterdam, 1989, pp. 239-253.

Implicit Rating and Filtering

David M. Nichols

Computing Department, Lancaster University,
Lancaster, LA1 4YR, UK

dmn@comp.lancs.ac.uk

Abstract

Social filtering systems that use explicit ratings require a large number of ratings to remain viable. The effort involved for a user to rate a document may outweigh any benefit received, leading to a shortage of ratings. One approach to this problem is to use implicit ratings: where user actions are recorded and a rating is inferred from the recorded data. This paper discusses the costs and benefits of using implicit ratings for information filtering applications.

Introduction

The increasing availability of information in computer-readable form is changing the nature of information searching. The users of information retrieval (IR) systems are faced with two problems: the sheer number of documents and a greater variation in the quality of those documents. The increasing heterogeneity of documents (both in quality, form and media) means that there is a greater need than ever before for tools to aid users in filtering and selecting relevant documents.

Malone *et al*. (1987) describe three forms of information filtering: cognitive (or content), economic and social. Content-based filtering is dominant in IR (e.g. Foltz and Dumais (1992)) – typified by profiles based on keywords. Economic filtering will become increasingly important as digital cash, micro-payments and secure payment technologies emerge from research laboratories onto the Internet. The third form, social filtering, has moved on from the original description (of the importance of the identity of the sender of a message) to several research projects and a few actively-used systems. The social filtering these systems perform is largely based on explicit ratings – where users rate a document on a pre-defined scale.

The rating of resources to enable collaborative (or social) filtering poses several problems: use of appropriate scales, motivation and incentives for evaluators (Avery and Zeckhauser, 1997), biased evaluators (Palme, 1997), avoiding the free-riding problem, achieving a critical mass of users (Oard and Marchionini, 1996) etc. Several of these problems are related to the explicit rating of items.

A small amount of other work has been done on using implicit information (Oard and Marchionini, 1996) - where ratings are automatically inferred from a user's behaviour. This paper will discuss the potential and the problems with using such implicit sources as a basis for filtering and recommending. The evidence for the use of implicit ratings is reviewed and the various types of implicit data available to digital library systems is described.

Implicit and Explicit Ratings

The use of explicit ratings is common in everyday life; ranging from grading students' work to assessing competing consumer goods (see Alton-Scheidl *et al.* (1997) for a review). Although some forms of rating are made in free text form (e.g. book reviews) it is frequently the case that ratings are made on an agreed discrete scale (e.g. star ratings for restaurants, marks out of ten for films etc). Ratings made on these scales allow these judgements to be processed statistically to provide averages, ranges, distributions etc. Implementations of ratings for computerised systems have largely followed this explicit approach.

A central feature of explicit ratings is that the evaluator has to examine the item and assign it a value on the rating scale. This imposes a cognitive cost on the evaluator – this is not necessarily a bad thing; society expects our teachers to think about the grades they give to their students. The value of many forms of rating derives from this intellectual effort and provides the justification for the remuneration that accompanies many rated information streams.

> Expert annotations require effort and have economic value, so the marketplace will undoubtedly assign them a price.
>
> (Oard and Marchionini, 1996)

When explicit ratings are used in social filtering systems (where the ratings of other users are used to generate predictions) the costs and benefits are clearly represented at the interface. The act of rating alters a user's behaviour from their normal pattern of reading - similarly, the choice of which items to examine is altered by providing a rated list. Moreover the benefits of any individual user's ratings are experienced by the other users of the system. This separation of costs and benefits has been noted as being very important in the failure of social computing systems (Grudin, 1994). Unless the user perceives some benefit for participating in the system then they have an incentive for leaving. Even worse, if the link between rating and receiving rated items is not reinforced then users may have an incentive to cease rating but continue to read. In such a system this could result in a lack of any ratings at all (Avery and Zeckhauser, 1997).

The problems for social filtering systems in acquiring explicit ratings have led to speculation that implicit ratings (gathered from user behaviour) may be a solution:

> We believe an ideal solution is to improve the user interface to acquire implicit ratings by watching user behaviors. Implicit ratings include measures of interest such as whether the user read an article and, if so, how much time the user spent reading it.
>
> (Konstan *et al.*, 1997)

The main motivation for using implicit ratings is that it removes the cost to the evaluator of examining and rating the item. Whilst there remains a computational cost in storing and processing the implicit rating data this can be hidden from the user. In a networked environment it is usually difficult for the user to separate network latency from extra application processing. Although there are clearly limits to user tolerances the storage/transport of implicit data at the client end is not a computationally intensive task.

As one of the main problems with obtaining explicit ratings is seen to be the acquisition costs (Oard and Marchionini, 1996) there should be a greater number of implicit ratings. Potentially, every user interaction with a system will generate implicit data – in fact we could move to a situation with too much data rather than the sparseness encountered by explicit rating approaches. Each implicit rating will probably contain less 'value' than an explicit rating but the appropriate cost-benefit trade-off for different types of implicit data will have to be determined empirically.

Acquiring Implicit Ratings

There are several types of implicit data that can, in principle, be captured and studied. (Stevens, 1992) uses three types of implicit data: read/ignored, saved/deleted and replied/not replied. (Morita and Shinoda, 1994) use reading duration in place of the read/ignore attribute. Table 1 shows the result of combining these forms with the types of usage data described in (Nichols, Twidale and Paice, 1997).

Action	**Notes**
Purchase (Price)	buys item
Assess	evaluates or recommends
Repeated Use (Number)	e.g. multiple check out stamps
Save / Print	saves document to personal storage
Delete	deletes an item
Refer	cites or otherwise refers to item
Reply (Time)	replies to item
Mark	add to a 'marked' or 'interesting' list
Examine / Read (Time)	looks at whole item
Consider (Time)	looks at abstract
Glimpse	sees title / surrogate in list
Associate	returns in search but never glimpses
Query	association of terms from queries

Table 1 Potential types of implicit rating information

Some of the data sources have additional information (e.g. a *Purchase* action has an associated *Price*) - these are indicated in parentheses. The actions are listed in an approximate ordering reflecting the importance of the type of data; it seems reasonable to conclude more from the purchase of an item rather than a simple inspection.

As Digital Libraries (DLs) and the Internet in general become a more commercial environment information providers will increasingly have *Purchase* information available. Elements of this style of investigation into users' purchase patterns are already being undertaken by business who provide 'loyalty cards'. Alongside the ostensible benefits to the customer the supermarket, for instance, gains data about the types and combinations of goods bought by consumers. These patterns can be used to inform marketing activities, e.g. at least one UK supermarket generates money-off vouchers for complementary and substitute goods at the checkout based on the type of goods bought. The extra information available from loyalty cards (or lifetime user IDs in a DL) can only reinforce this trend.

Although we are discussing implicit data it is also possible to gather implicit data from an explicit rating scenario. The *Assess* category distinguishes those events when a evaluator chooses *not* to rate an item when they could have done so. Hence this category would not contain any reference to the actual value of a rating only the fact that a rating had, or had not, occurred.

The *Repeated Use* category in Table 1 could really refer to any of the other categories of data. However, it has an appealing analogue with conventional library practice, that of date stamps in the back of a borrowed book. Items that a user wishes to preserve for some purpose are often *Saved* to personal file space or *Printed.* The *Delete* category will clearly only apply to certain types of information stream (e.g. Usenet News) and differs from the others in that it expresses a negative judgement.

The field of Library & Information Science (LIS) has examined the use of citations in considerable depth. The *Refer* category contains all those instances where one information item links to another item; this includes traditional academic citations as well as less formal links such as hyperlinks on Web pages or the threaded links between Usenet News articles. In some interactive information environments (e.g. Usenet) users can *Reply* to items they encounter; either back to the sender or via a public forum. The *Time* taken to compose this reply may also be available. In many environments a user will *Mark* certain items as being of particular interest so that they can easily return to them, e.g. Web browsers enable hotlists or bookmarks to be recorded.

The next three categories, *Examine*, *Consider* and *Glimpse*, all refer to the same action: the user reading a document (or document surrogate). Systems usually allow users to read a shortened or summary version of a document; bibliographic databases often have an abstract rather than the full-text of an article.

At the bottom of the list in Table 1 the action *Associate* refers to items which are closely connected to those that are examined, e.g. items in the second page of hits which is never reached by the user. The action *Query* refers to query terms which have been used by searchers and can then be reused by subsequent searchers who use related terms (Koenig, 1990).

The collection of these types of implicit data does not pose difficult technological problems: many information access tools could easily be modified to record most of the categories of data in Table 1. In addition, there is a considerable body of research in LIS on the closely related field of transaction log analysis, e.g. (Flaherty, 1993). Data acquired through transaction log analysis has been passed back to designers to refine their systems (typically through user interface modifications). In contrast, an implicit rating system directly accesses the data to modify the system – there is no human in the feedback loop.

Implicit Rating Systems

There appears to have been little work done on implicit ratings, a recent survey (Oard and Marchionini, 1996) mentions only two sources: Morita & Shinoda and Stevens. There are two other major sources: the PHOAKS system (Hill and Terveen, 1996) and GroupLens (Konstan, *et al.*., 1997). The GroupLens project have reported the most interesting results with regard to time-based implicit data; they summarise the situation as:

> Our initial studies show that we can obtain substantially more ratings by using implicit ratings and that predictions based on time spent reading are nearly as accurate as predictions based on explicit numerical ratings. … Our results also provide large-scale confirmation of the work of Morita and Shinoda in finding the relationship between time and rating holds true without regard for the length of the article.
>
> (Konstan, *et al.*., 1997)

Both GroupLens and Morita & Shinoda judge regard time spent reading as a good candidate as a basis for filtering. There is however a difference between the two sets of experiments - the GroupLens data is derived from an explicit rating scenario whereas Morita & Shinoda use post-session rating.

The GroupLens experimental model (Model 1) is:

1. users evaluate news items
2. system collects implicit data
3. compare the explicit and implicit data

From these results they show the similarity of the two sources. The Morita & Shinoda model (Model 2) is:

1. users read news items
2. system collects implicit data
3. system predicts using implicit data
4. users return to items and evaluate
5. compare the explicit and implicit data

Neither model is perfect, in Model 1 it may be premature to use implicit data derived and verified from an explicit rating scenario. User behaviour may be significantly different when they are reading 'normally' - consequently the correlation between implicit data and user judgements may not be as strong or as reliable. In Model 2, the users make their evaluations on their second view of the items - when they have already seen the rest of the items. It seems likely that viewing the other items will alter their assessment of the earlier items.

Indeed, (Oard and Marchionini, 1996) make the general point as to how we can generalise these experimental results to real-world settings where users are distracted and interrupted. The test subjects in the Morita & Shinoda experiment were asked to read items continuously, a very different scenario from their usual news-reading habits.

A separate source of implicit data for Web users are the files of bookmarks, or hotlists, of Web pages; the *Mark* category from Table 1. The Siteseer system (Rucker and Polanco, 1997) uses the overlap between bookmark files to create *virtual communities of interest* and then recommends URLs pages from a users' *virtual neighbours*. The Group Asynchronous Browsing (Wittenburg *et al.*, 1995) uses a similar approach but aims to create an enhanced browsing structure rather than an explicit recommendation.

A simple example of using *Purchase*-based implicit data is currently in operation at Amazon.com (Amazon.com, 1997), the online bookstore. When the entry for a book is displayed other titles bought in conjunction with it are also shown, e.g. *The Design of Everyday Things* by Donald Norman produces:

Check out these titles! Readers who bought The Design of Everyday Things also bought:

- The Visual Display of Quantitative Information; Edward R. Tufte
- Visual Explanations : Images and Quantities, Evidence and Narrative; Edward R. Tufte
- Things That Make Us Smart : Defending Human Attributes in the Age of the Machine; Donald A. Norman

http://www.amazon.com/exec/obidos/ISBN=0385267746/5932-9389921-467955

The other major project using implicit data is PHOAKS; this system scans Usenet postings to find mentions of URLs which it takes as an implicit recommendation. This is an example of the *Refer* category from Table 1. The PHOAKS system has several heuristics to try to eliminate URLs that contain little value, e.g. those contained in signatures. It seems likely that similar 'pruning' techniques will be necessary to use implicit data efficiently.

Rating Scenarios

	Give Explicit Ratings	**Give Implicit Ratings**	**Receive Predictions**	**Examples**
1	–	–	–	normal Usenet reading
2	–	–	_	freeloader, client
3	_	–	–	rating service
4	_	_	–	rating service
5	_	_	_	GroupLens
6	_	–	_	GroupLens
7	–	_	–	implicit data provider only
8	–	_	_	implicit data provider only

Table 2 Scenarios for implicit and explicit rating

Table 2 shows the possible scenarios involving a combination of implicit rating, explicit rating and receiving predictions. Case 1 is the present situation, where most readers do not use ratings and do not receive predictions. In case 2 the user receives the benefit of predictions but does not contribute any ratings; such a user could be a freeloader or a client of a rating service – depending on whether they pay for the predictions. Cases 3 and 4 could describe the behaviour of such a rating service – where predictions are not important.

Cases 5 and 6 describe the users of a social filtering system such as GroupLens – giving ratings and receiving predictions. Case 7 describes the situation where a user allows their implicit data to be gathered but does not receive any predictions. Case 8 describes the scenario where the user does receive those predictions.

Any social filtering system will have users who can be located within these different scenarios but a successful system will have to maintain appropriate ratios between their users. A system with too many freeloaders from case 2 will soon cease to be viable.

Conclusion

The limited evidence available suggests that implicit ratings have great potential but their effectiveness remains unproven. As with many technologies implicit rating may first be combined with existing rating systems to form a hybrid system. One approach is to use implicit data as a check on explicit ratings, e.g. if an evaluator is explicitly rating an item then there should be some corresponding implicit data to confirm that she has actually examined it. If there is no evidence to suggest that the evaluator has examined an item then perhaps their rating should be ignored, or reduced in importance. Conversely, an evaluation with a relatively long 'examine time' may be increased in importance.

The existing systems which capture implicit data (such as Web servers) have generated some concerns amongst the general population of users about privacy. Although we can discuss the possibilities of using implicit data – systems need to be 'socially' accepted in order to be successful. This is especially true of social filtering systems - whose very power comes from a wide take-up of different users.

Acknowledgements

I would like to thank Mike Twidale, Paul Rayson and Chris Paice for stimulating discussions about the topics in this paper.

References

Alton-Scheidl, R., Schumutzer, R., Sint, P.P. and Tscherteu, G. (1997), Voting and rating in Web4Groups, in Alton-Scheidl, R., Schumutzer, R., Sint, P.P. and Tscherteu, G. (Eds.), *Voting, Rating, Annotation: Web4Groups and other projects: approaches and first experiences*, Vienna, Austria: Oldenbourg, 13-103.

Amazon.com (1997), *Amazon.com*, http://www.amazon.com,

Avery, C. and Zeckhauser, R. (1997), Recommender systems for evaluating computer messages, *Communications of the ACM*, 40(3), 88-89.

Flaherty, P. (1993), Transaction logging systems: a descriptive summary, *Library Hi Tech*, 11(2), 67-78.

Foltz, P.W. and Dumais, S.T. (1992), Personalized information delivery: an analysis of information filtering methods, *Communications of the ACM*, 35(12), 51-60.

Grudin, J. (1994), Groupware and social dynamics: eight challenges for developers, *Communications of the ACM*, 37(1), 92-105.

Hill, W. and Terveen, L. (1996), Using frequency-of-mention in public conversations for social filtering, *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW'96)*, Cambridge, MA, ACM Press, 106-12.

Koenig, M.E.D. (1990), Linking library users: a culture change in librarianship, *American Libraries*, 21(9), 844-9.

Konstan, J.A., Miller, B.N., Maltz, D., Herlocker, J.L., Gordon, L.R. and Riedl, J. (1997), Applying collaborative filtering to usenet news, *Communications of the ACM*, 40(3), 77-87.

Malone, T.W., Grant, K.R., Turbak, F.A., Brobst, S.A. and Cohen, M.D. (1987), Intelligent information sharing systems, *Communications of the ACM*, 30(5), 390-402.

Morita, M. and Shinoda, Y. (1994), Information filtering based on user behaviour analysis and best match text retrieval, *Proceedings of the 17th ACM Annual International Conference on Research and Development in Information Retrieval (SIGIR'94)*, Dublin, Ireland, Springer-Verlag, 272-81.

Nichols, D.M., Twidale, M.B. and Paice, C.D. (1997), *Recommendation and Usage in the Digital Library*, Technical Report CSEG/2/97, Computing Department, Lancaster University.

Oard, D.W. and Marchionini, G. (1996), *A Conceptual Framework for Text Filtering*, Technical Report CAR-TR-830, Human Computer Interaction Laboratory, University of Maryland at College Park.

Palme, J. (1997), Choices in the Implementation of Rating, in Alton-Scheidl, R., Schumutzer, R., Sint, P.P. and Tscherteu, G. (Eds.), *Voting, Rating, Annotation: Web4Groups and other projects: approaches and first experiences*, Vienna, Austria: Oldenbourg, 147-62.

Rucker, J. and Polanco, M.J. (1997), Personalized navigation for the web, *Communications of the ACM*, 40(3), 73-5.

Stevens, C. (1992), *Knowledge-Based Assistance for Accessing Large, Poorly Structured Information Spaces*, PhD Thesis, Department of Computer Science, University of Colorado.

Wittenburg, K., Das, D., Hill, W.C. and Stead, L. (1995), Group asynchronous browsing on the World Wide Web, *Proceedings of the Fourth International World Wide Web Conference*, Boston, MA, O'Reilly & Associates, 51-62.

Choices in the Implementation of Rating

By Jacob Palme, first version January 1997, revised version July 1997.

This document is available on the WWW on URL
http://www.dsv.su.se/~jpalme/select/rating-choices.html

Abstract

This paper discusses how an Internet-based collaborative filtering system can be implemented and presents references to descriptions of some existing such systems. The paper discusses who should input ratings, how ratings can be used, and presents an architecture for a rating and filtering system. This architecture is defined such that different people at different places can implement different modules in the architecture.

Table of contents

What is Rating?

By rating is meant services by which your selection of resources to read is guided by the quality of the resources, as specified by people who have read the resource. Rating is also known under the terms “collaborative filtering” or “social filtering”.

In the Internet, rating may be applied to many kinds of resources, like web pages, messages, electronic journal papers, public domain software.

The purpose of rating may be to increase the quality of the resources you read, or to avoid certain resources deemed unsuitable in certain communities for certain groups of readers (example: violence, pornography).

In the world before the Internet, rating was commonly provided by services such as:

- Newspapers, magazines, books, which are rated by their editors or publishers, selecting information which they think their readers will want.
- Consumer organisations and trade magazines which evaluate and rate products.
- Published reviews of books, music, theatre, films, etc.
- Peer review method of selecting submissions to scientific journals.

Rating is further described in [3].

Vocabulary

This vocabulary is partly based on [1]:

Category

Value system used in rating, example “1;2;3” or “objectionable; acceptable”. Also known as “dimension”.

Censorship

See *Parental control.*

Content label

A data structure indicating a rating of a particular resource or set of resources. Also known as “rating” or “content rating”.

Label bureau

A computer system that supplies, via a computer network, ratings of resources. It may or may not also provide the resources themselves.

Parental control

Software and services for use by parents and teachers to control children's' usage of the Internet. The main goal of such software is to make it impossible without special privileges do download forbidden information. Such systems might thus also be labelled *Censorship systems*. Compare with *Peer collaborative filtering*.

Peer collaborative filtering

Collaborative filtering systems to be used among peers to aid each other in finding the most interesting information. Compare with *Parental control.*

PICS

Platform for Internet Content Selection, a W3C specification for format and protocols for rating.

Rating service

An individual or organisation that assigns labels according to some rating system, and then distributes them, perhaps via a label bureau or via CD-ROM.

Rating system

A method for rating information, consisting of one or more categories.

Resource

Object or document on the net which can be rated, such as web page, newsgroup article or downloadable software.

Scale

The range of permissible values for a category.

Existing rating software and services

Peer Collaborative Filtering Versus Parental Control

Most rating software and services at present (summer 1997) are defined for the specific goal of protecting children from information which is regarded as unsuitable to them. This is thus a kind of censorship system, meant to be used by parents and teachers to control children's' usage of the Internet.

The software in such systems will partly work differently than collaborative filtering to be used among peers to aid each other in finding the most interesting information. Parental control software will make it impossible, without special privileges, do download forbidden information. Peer collaborative filtering software, on the other hand, aims at giving information to the user, and need not always remove or stop less desirable information. Also, the categories and scales are different. Typical categories in parental control is violence, sex, nudity, language, or age at which children should be allowed to see this information. Typical categories in peer rating systems might be quality or newsworthiness.

Parental control

The PICS standard [1], [2], [4] was mainly developed for parental control, and most existing implementations of PICS have this goal, even though the PICS protocols are equally useful for peer collaborative filtering.

Many systems and services for parental control are available, such as Bess, Cyber Patrol, CyberSitter, Cyber Snoop, Gulliver's Guardian, Net Nanny, NetShepherd, etc. An overview with links to further information on such systems can be found at [5].

Links to parental control systems using the PICS standard can be found at [6].

One very well-known such services is the Recreational Software Advisory Council for the Internet (RSACi) [9]. The basis of RSACi is to give objective descriptive information about rated resources, not subjective judgements. The intention is that this would enable the owner of a resource to rate his own resources. RSACi rates resources on four dimensions: violence, nudity, sex and language. A questionnaire is provided with simple yes/no questions. By answering this questionnaire, RSACi ratings are automatically produced.

The main alternative to RSACi are systems and services based on subjective judgement of what is suitable and not suitable for children of a certain age. Such services typically provide an age level, saying that a certain resource is not suitable for children below this age level. The most well-known such system is the Motion Picture Association of America (MPAA) system [7], [8].

Peer Collaborative Filtering Systems and Services

Some wellknown collaborative filtering systems and services at present (summer 1997) are:

Firefly (http://www.firefly.net) is a company which both sells collaborative filtering software and services. Firefly is used by other Internet service providers, for example Yahoo claims to provide a special ratings-based service MyYahoo (http://my.yahoo.com).

A description of how Firefly collaborative filtering works can be found at [10]. Firefly say that they compute correlations between the scores given to resources by different users, and finds those other users whose score has highest correlation to your scores. Resources which they rate highly are then suggested as of interest to you. Firefly further says that they are using a system called Feature-Guided Automated Collaborative Filtering. This means that the information space is divided into different subject areas, and collaborative filtering is then performed only within such an area.

Net Shepherd (http://www.shepherd.net) started as a parental control service, but has evolved into the area of peer collaborative filtering. The description of their service in [11] seems to indicate that they (summer 1997) are only providing majority ratings by all raters, not individually selected ratings from people with similar interests and values as themselves.

Net Perceptions(http://www.netperceptions.com) markets a collaborative filtering system called GroupLens [13]. GroupLens can collect explicit ratings, or can implicitly estimate ratings based on the time a user uses to view a resource. It is mainly marketed for organisations who want to provide collaborative filtering to their own users, and is not marketed as a global collaborative filtering systems for resources all over the Internet. GroupLens was originally developed at the MIT Centre for Coordination Science [14].

Sepia Technologies, Inc. in Quebec, Canada, has developed a collaborative filtering system for movies, music and books [15].

The PICS standard

The PICS standard, developed by the World Wide Web Consortium [1], [2], [4] is a very general-purpose standard for supplying ratings. Within the PICS standard, it is possible to define your own rating system, with your own categories and scales. Your rating system can contain several different categories with different scales. For example, the four RSAC scales of violence, nudity, sex and language can as easily be accommodated as the MPAA scales of age limits for children.

When you use PICS, you first define your rating categories and scales and specify these in a particular notation [1]. Here is an example of a description of a category in a rating system specification:

```
(category
   (transmit-as "hue")
   (label (name "blue")  (value 0))
   (label (name "red")   (value 1))
   (label (name "green") (value 2)))
```

When a rating system has been defined, it is then possible to distribute rating labels [2]. A rating label contains a description of one or a set of resources. It is possible to define a rating label for a whole web site, but then to supply different rating labels to subspaces within that web site or to individual resources. The rating for the whole web site is then only used when no more narrow rating label is available for a particular resource.

For HTML resources, the rating labels can be put as META fields in the HEAD of the HTML text, so that it is downloaded as part of the resource. PICS also specifies protocols for a web site to provide a special server for providing ratings of its web pages, and protocols for services which provide ratings also for other web pages than its own.

The resource being rated is identified by its URL. Since URLs [12] are not only available for web pages, but also for e-mail messages, Usenet News newsgroups and messages, etc., PICS can be used to rate all resources for which URLs are defined.

Some Problems with Rating

Some problems which can cause rating to work less well are:

1. Too few ratings are provided to provide a good basis for rating.
2. It may be difficult to collect ratings from users. Some systems solve this by implicitly guessing user ratings from the time the user spends reading a resource.
3. Some raters may not do a good work of rating.
4. People can unduly influence the rating to favour their own work, or work by their friends, relatives or co-workers.

5. Ratings may not be set by people with the same values and views at yourself. For example, an expert in an area may prefer other choices than beginners. A resource which experts give bad ratings to, may be good for beginners. Also your values may influence your choices, for example political values may influence whether you prefer analysises based on a conservative, liberal or class struggle viewpoint, or a religious person may have different preferences than a cynical/sophisticated “modern” person.

Design of rating systems which better handle one of the above requirements may be less good for other requirements. For example, restricted selecting of who may provide the ratings may give higher-quality ratings (at least if your values and views are the same as of those providing the rating) but reduce the amount of ratings and rated resources available.

Choices for Rating

The table below discusses the interaction of two choices in rating system design.

The horizontal axis represent the choice of restricting peoples' rights to submit ratings, the vertical axis represents the choice of whose ratings to use for your selection needs.

Table 1: Whose ratings are used where?

		Right to rate a resource	
		Everyone can input any rating (except limitations that you cannot rate your own or your friends' resources)	The right to input ratings is limited in some other way, to select people most proficient at providing good ratings in some way
Use of ratings in filtering	An average of all ratings set by everyone or by members of your peer group.	Advantage: Lots of ratings available. Disadvantage: Ratings may not agree with your personal preferences.	Advantage: Better rating, may avoid misuse. Disadvantage: May reduce the amount of ratings available.
	Ratings of people with similar views to yourself are preferably used through an automatic mechanism of comparing your ratings with those of other people.	Complex to implement, but might provide very good ratings for your views and requirements. Also, this might give larger availability of ratings, since only by giving your own ratings on resources can your preferences be matched to those of other people.	This combines two different ways of trying to achieve the same thing: Ratings set by those providing good ratings are given priority. This combination should not be used unless carefully analysed, since otherwise the two services can interact in unsuitable ways.

To select only certain people who are allowed to provide ratings, or to let anyone provide ratings, but base your selections on ratings made by people with your values and views, are two alternative methods of getting higher-quality ratings. Is it an advantage to combine both methods, or will they interact so that one method is better than the other?

Resources to be rated

Common goals of rating:

- Messages sent via mailing lists and other e-mail messages.
- Articles in Usenet News and messages in other group communication systems.
- Articles in the many journals and magazines which are published on the web.
- Web pages containing scientific papers.
- Web pages containing popular science, art, etc.
- Other kinds of web pages, Gopher documents, FTP documents.
- Public domain and share-ware software.
- Articles in Usenet News and contributions in other conference systems like Web4Groups.
- Sets of resources, such as web sites, or subareas within a web site.

A single common rating for a set of more than one resource (such as a site or all resources with a certain initial part of there URLs) has both pros and cons.

Pro: It is less effort to rate sets than every single resource, which means that more ratings will be available.

Con: The quality may vary between resources within the same set.

To reduce the disadvantage, rating on sets of resources should not encompass the whole of heterogeneous web sites. As an example, a university department should sometimes be rated separately for different researchers or research groups within the department.

Rating systems

Rating services may use different rating systems. A rating system to avoid objectionable resources may for example use terms like "unsuitable for children below 15 years" or "nakedness" while a rating system for movies may use a system of * to *****.

Suggested rating system for rating of other people's resources

A category scale from 1 to 10, defined as follows:

1. Of no value at all, to be avoided.
2. Of very little value.
3. Of little value.
4. Maybe of some value.
5. Of some interest.
6. Of interest, but not essential.
7. Very interesting and/or valuable.
8. Highly interesting and valuable.
9. Close to excellent.
10. Excellent.

Suggested rating system for rating of your own resources

Note: These categories use terms which are not easy to misuse to give your own resources too high ratings:

1. Flaming, jokes, advertisements, non-serious items.
2. Ordinary personal viewpoint or discussion item.
3. Very well-considered personal viewpoint or discussion item.
4. Poems, short stories, novels.
5. Art, music, fictional videos, etc.
6. Well-considered and researched monograph.
7. Article published in edited journal, book published by book publishing company of the kind which publishes quality books.
8. Masters thesis at a university or of comparable quality.
9. Paper accepted for publication in peer-reviewed scientific journal.
10. Doctoral thesis or of comparable quality.

Architecture of a rating system

Source

This architecture was developed for and part of the proposal for an EU grant to a research project on intelligent and collaborative filtering with the name SELECT. This proposal has been recommended for acceptance by the EU, and the research project is expected to start in January 1998.

Modularisation of a filtering and rating system

If a rating and filtering systems is to be implemented by people and organisations in many different countries, then the rating and filtering system need be split into well-defined modules with a well-defined interface between them. Here is a first attempt to define this set of modules:

Figure 1: Relations between modules

(Arrows indicate the direction of information flow, not the direction of control)

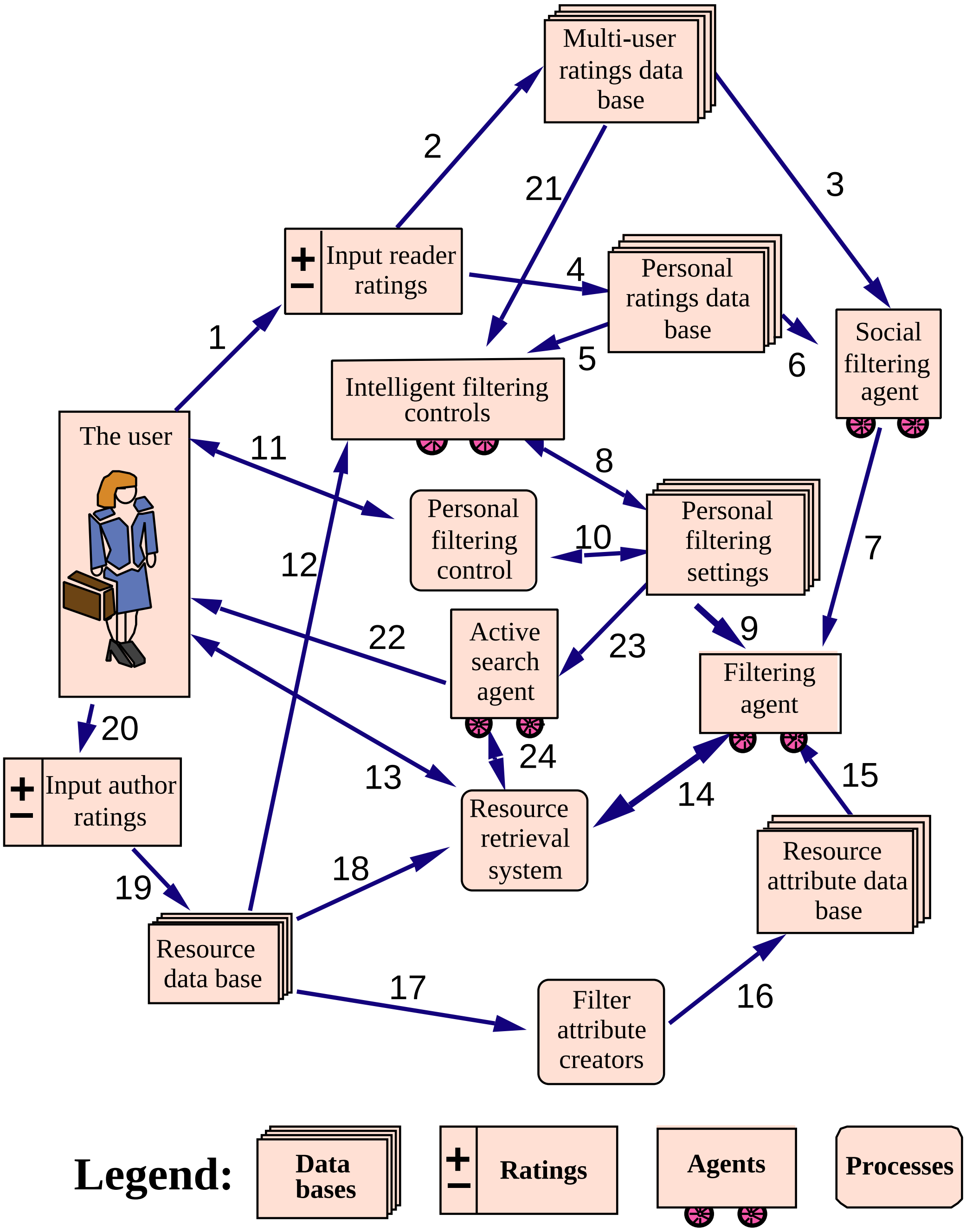

Multi-user ratings data base
2
21
3
Input reader ratings
4
Personal ratings data base
1
5
6
Social filtering agent
Intelligent filtering controls
The user
11
8
Personal filtering control
10
Personal filtering settings
7
12
22
Active search agent
23
9
Filtering agent
20
24
Input author ratings
13
14
15
Resource retrieval system
19
18
Resource attribute data base
Resource data base
17
16
Filter attribute creators
Legend:
Data bases
Ratings
Agents
Processes

Table 2: Modules in the system

Name	Description	Relations to other modules
Input of author ratings	An author can give his own resources ratings, using the scale above for author-specified ratings.	Input from **user interface** (20), stored in **RFC822** or **HTML header** (19), retrieved with the resource itself.
Input of reader ratings	A reader can, when reading an article, a message or a web resource, specify a rating using the scale above for reader-specified ratings.	1. Input from **user interface** (1). 2. Ratings are moved to a **personal ratings data base** (4), which can be used to automatically deduct better intelligent filtering methods for this user, and also: 3. Ratings are moved to a **multi-user ratings data base** (2), to aid other people's filtering.
Personal ratings data base	A data base, accessible only by a certain person and agents working for that person. The data base contains a list of messages and ratings. The data base should have news control, so that an agent connecting to this data base can download the new ratings put into the data base since the last time this agent connected to this data base.	**Intelligent filtering controls** (5) can scan this data base, and deduct filtering conditions based on its contents. **Social filtering agents** (6) can match the personal choices in this data base with the personal choices of other people, found in a multi-user ratings data base, to deduce which other persons have similar preferences to this user, so that their ratings can be used to guide this user.

Multi-user ratings data base	A data base, accessible to rating and filtering agents. The data base contains a list of messages and ratings. For every rating, the data base contains a uni-directional encryption of the e-mail-address of the person who provided this rating. In this way, it is possible to identify ratings made by the same person, without knowing who this person is. The data base should have news control, so that an agent connecting to this data base can download the new ratings put into the data base since the last time this agent connected to this data base.	Used by and accessible to different kinds of agents like **social filtering agents** (3) and **intelligent filtering controls** (21). Can also be used as a research data base for development of better ratings and filtering systems, and should thus be accessible for researchers. To avoid misuse, it should maybe not be accessible to anyone using any kind of software (since there is a risk of deriving the real user ID from the encrypted user ID).
Filter attribute creators	A filter attribute creator is a piece of software which derives filter attributes from a resource. Basic attributes are words (very common words excluded). Words may be transformed to a canonical form and be extended with synonyms. Other attributes are length of original and of quoted text, percentage of multi-syllable words and other genre-indicators, use of graphics and advanced HTML constructs, etc.	Takes as input resources (**articles, messages** and **web pages** (17)) and produces additional data which is stored in a **resource attribute data base** (16).

Resource attribute data base	A data base of attributes for a resource. The attributes may be stored in inverted form, so that you can rapidly search for resources with certain attributes or attribute combinations (this is often done by network search engines like Alta Vista or Euroseek).	Input from **filter attribute creators** (16). Output to **filtering and searching agents** (15).
Intelligent filtering controls	Agent which reads the **Personal ratings data base**, looks at the resources you liked and disliked, and deduces filtering conditions to find the resources you like and not those you dislike. Note that this agent does not perform the actual filtering, it just provides input to the **Personal filtering settings**, which are then used to control the actual filtering..	Input from **Personal ratings data base** (5). Output to **Personal filtering settings** (8).
Personal filtering settings	Settings which controls your filtering agents. These settings include code in a of language for specifying filtering conditions, probably based on Boolean algebra.	Input from **Intelligent filtering controls** (8), output to **Filtering agents** (9) and input and output from **Personal filtering control** (10).
Personal filtering control	Lets you see and modify your personal filtering settings.	**User interface** (11) and input and output to **Personal filtering settings** (10).

Filtering agent	Agent which uses the personal filtering settings to perform filtering of resources for you.	Input from **Personal filtering settings** (9) and from **Resource attribute data base** (15), and from the **Resource retrieval system** (14).
Social filtering agent	Agent which uses the social filtering information to perform filtering of resources for you.	Used as a subsystem by your **Filtering agent** (7), uses data from **Multi-user ratings data base** (3) and **Personal ratings data base** (6).
Resource retrieval system	System for getting resources from the Internet. Examples of such systems: E-mail system, Usenet News system, Web browser, Web search index provider, Web4Groups system.	Input and output from **user interface** (13), and input and output from **Filtering agent** (14).
Resource data base	Existing data bases of Internet resources, such as part or whole of the WWW information space, mailing list archives or Usenet news servers.	An author can **Input author ratings** (1) of the resources he has authored, for example, for HTML documents, such ratings can be stored as META fields in the HEAD.
Active search agent	Agent which automatically scans the net, searching for information of interest to a particular user.	Controlled by **Personal filtering settings** (23), scans the net (**Resource retrieval system**) (24) and delivers results to the user (22).

Table 3: Interfaces to be defined

No.	Related modules	Operation	Format	Protocol
1	**User** and **Input reader ratings**.	User interface.	To be defined by user interface experts.	HTML/HTTP.

2	**Input reader ratings** and **Multi-user ratings data base.**	**Input reader ratings** stores ratings in the **Multi-user ratings data base**.	Might be based on PICS. PICS may have to be extended with a method of transmitting the name of the rater?	To be defined, probably as a variant of HTTP.
3	**Multi-user ratings data base** and **the Social filtering agent.**	**The Social filtering agent** can retrieve information from the **Multi-user ratings data base.**	To be defined.	To be defined, probably as a variant of HTTP. We have to decide whether much information is transported to the **Social filtering agent**, or whether the main processing is done in the Multi-user ratings data base and only the results transported to the **Social filtering agent**.
4	**Input reader ratings** and **Personal ratings data base.**	**Input reader ratings** stores ratings in the Personal **ratings data base**.	Can possibly be similar to 2 above.	Can possibly be similar to 2 above.
5, 21	**Personal ratings data base**, **Multi-user ratings data base** and the **Intelligent filtering controls.**	**The Intelligent filtering controls** can retrieve information from the **Personal ratings data base.**	Can possibly be similar to 3 above. but the intelligent filtering controls may need more information.	Can possibly be similar to 3 above but the intelligent filtering controls may need more information.

6	**Personal ratings data base** and the **Social filtering agent.**	**The Social filtering agent** can retrieve information from the **Personal ratings data base.**	Can possibly be similar to 3 above.	Can possibly be similar to 3 above.
7	**Social filtering agent** and **filtering agent**	The **Social filtering agent** is used as a subsystem by the **Filtering agent**.	Can PICS be used?	Can PICS be used?
8	**Intelligent filtering controls** and **Personal filtering settings.**	The **Intelligent filtering controls** can modify the **Personal filtering settings**.	Format for personal filtering settings is needed. Might be based on Boolean algebra, but we should also look at fuzzy logic. We should also look at Compassware (http:/www.compassware.com).	To be defined, probably as a variant of HTTP.
9	**Personal filtering settings** and **Filtering agent.**	The **Filtering agent** can retrieve the **Personal filtering settings**.	See 8.	To be defined, probably as a variant of HTTP.
10	**Personal filtering control** and **Personal filtering settings.**	The **Personal filtering control** can retrieve and modify the **Personal filtering settings**.	See 8.	To be defined, probably as a variant of HTTP.

11	**User** and **Personal filtering control.**	User interface There should be a simple mode for people who do not want to learn the language for specifying filtering conditions, and an advanced mode for those who wants to learn this language.	To be defined by user interface experts.	HTML/HTTP.
12	The **Intelligent filtering controls** and the **Resource data base.**	The **Intelligent filtering controls** can retrieve resources from the **Resource data base**.	MIME resource formats.	HTTP, FTP, Gopher, NNTP, Web4Groups.
13	**User** and **Resource retrieval system.**	This is an augmented version of the normal user interface for the **Resource retrieval system**.	To be defined by user interface experts.	As used in the resource retrieval system.
14	**The Filtering Agent** and **the Resource retrieval system.**	The **Resource retrieval system** can enlist the help (input and output) from the **Filtering agent**.	To be defined.	To be defined.
15	**Resource attribute data base** and **Filtering agent.**	The **Filtering agent** can retrieve attributes from the **Resource attribute data base**.	Variant of PICS?	To be defined, probably as a variant of HTTP.
16	**Resource attribute data base** and **Filter attribute creators.**	The **Filter attribute creators** stores its results in the **Resource attribute data base**.	Variant of PICS?	To be defined, probably as a variant of HTTP.

17	**Resource data base** and **Filter attribute creators.**	The **Filter attribute creators** use the normal access protocol to the **Resource data base** (such as HTTP, NNTP, POP, Web4Groups access protocol).	MIME resource formats.	HTTP, FTP, Gopher, NNTP, Web4Groups.
20	**User** and **Input author ratings.**	User interface.	To be defined by user interface experts.	HTML/HTTP.
21	**See 5 above**			
22	**Active agent** and **User**	The **Active agent** delivers its results to the user	To be defined by user interface experts.	This might be through the user interfaces already provided by one of the Resource retrieval systems used.
23	**Active agent** and **Personal filtering settings**	The **Personal filtering settings** are used by the user to guide the **Active agent**.	See 9 above.	See 9 above.

References

[1] Rating Services and Rating Systems (and Their Machine Readable Descriptions), by Jim Miller, Paul Resnick and David Singer, available at URL: http://www.w4.org/PICS/services.html. *This document, together with [2], is the official definition of the PICS standard.*

[2] PICS Label Distribution Label Syntax and Communication Protocols, by Jim Miller, Tim Krauskopf, Paul Resnick and Win Treese, URL http://www.w3.org/pub/PICS/labels.html. *This document, together with [1], is the official definition of the PICS standard.*

[3] Voting and Rating: Perspectives for Information Collection, Decision Making and

Collaborative Rating Using Web4Groups by Austrian Academy of Sciences. Internal Web4Groups paper, November 1996.

[4] PICS: Internet Access Controls Without Censorship, by Paul Resnick and James Miller, URL: http://www.bilkent.edu.tr/pub/WWW/PICS/iacwc.htm. *An introductory overview to PICS.*

[5] The Kids on the Web: Safety on the Net, by Brendan Kehoe, URL: http://www.zen.org/~brendan/kids-safe.html. *A list of links to different parental control systems and services.*

[6] Pics Third-Party Rating Services, URL: http://www.w3.org/pub/WWW/PICS/raters.htm. *A list of links to services based on the PICS standard.*

[7] The MPAA Rating Systems, URL: http://ficus-www.cs.ucla.edu/ficus-members/reiher/film_miscellany/ratings.html. *An introduction to the MPAA rating system.*

[8] The Voluntary Movie Rating system, by Jack Valenti, URL http://www.mpaa.org/ratings.html. *An overview of the MPAA rating system.*

[9] Recreational Software Advisory Council, URL: http://www.rsac.org/. *Home page for RSAC, with links to many informational documents on RSAC.*

[10] Collaborative Filtering Technology: An Overview. URL: http://www.firefly.net/products/CollaborativeFiltering.html. *A description of how the Firefly collaborative filtering system works.*

[11] Net Shepherd 2.0 Frequently Asked Questions. URL: http://www.shepherd.net/products/NetShepherd2.0/faqs.HTM. *A description of the peer collaborative filtering service provided by Net Shepherd.*

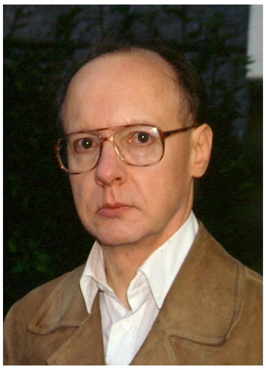
[12] T. Berners-Lee, L. Masinter, M. McCahill, "Uniform Resource Locators (URL)", by T. Berners-Lee, L. Masinter, M. McCahill , Internet RFC 1738, December 1994. *This is the Internet standard for URLs. There are numerous other IETF standards specifying the URL format for different kinds of resources.*

[13] Building Customer Loyalty and Profitable 1-to-1 Customer Relationships with Net Perception's GroupLens™ Recommendation Engine. URL: http://www.netperceptions.com/product_whitepaper.html. *A description of the collaborative filtering system from Net Perceptions.*

[14] GroupLens: An Open Architecture for Collaborative Filtering of Netnews, by P. Resnick et al. Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work, Chapel Hill, Pages 175-186, and at URL http://ccs.mit.edu/CCSWP165.html.

[15] Collaborative Filtering The SEPIA Suggestion Box ®, at URL http://www.sepia.com/suggestion_e.html.

About the author



Jacob Palme (http://www.dsv.su.se/~jpalme) is non-tenured professor of Computer Science at Stockholm University and the KTH Technical University. He has been doing both technical and social science research in the area of Computer Mediated Communication (CMC) since 1975. He has also participated in ISO and IETF standards work in the CMC area and is the author of numerous textbooks in Computer Science, the latest with the title *Electronic Mail* (http://www.dsv.su.se/~jpalme/e-mail-book/e-mail-book.html) was published in 1995.

Institutional Rating in Everyday Life

Peter Paul Sint, Austrian Academy of Sciences, Research Unit for Socio-Economics

Abstract

Rating on the Internet seems to be a rather new kind of activity. However, traditionally a quite large number of institutional rating mechanisms are established. We have studied a number of such rating methods to gain some insights about the role rating can play in the actions of of individuals and institutions.

Introduction

We all rate objects and events in our environment and our life. We evaluate persons, their actions towards us and towards others. We rate objects according to their usefulness for specific purposes and actions for their appropriateness. We all exchange personally assessments of those items: be it by gossip, by serious consultation or by formal channels.

If we try to introduce rating schemes in the electronic communication environment it could help to look at rating in traditional social environments. Our emphasis in this respect will be on formalised, institutionalised forms of rating in our social communities. Having set this goal, it became evident very soon that many different forms of ratings exist: Grading in schools, rating of personnel in companies, assessment and evaluation of projects and programmes, rating of consumer products by consumer associations, professional evaluation of working tools, assessment of performance in sports, awarding prices, medals and honours in literature and arts.

What are the common features of those schemes? What can we learn from them in setting up new rating mechanisms? Will rating mechanisms in the future be modelled after traditional forms? Or will there emerge new forms? What are the ingredients to be preserved? In what way do they fail? What additional benefits has electronic communication to offer?

Studying rating procedures, we have to take into account several different features:

What are the functions of rating mechanisms? What purposes does a rating fulfil? What are its uses for the recipient?

How is rating done? What are the methods of rating? Experts, panel, formal mathematical methods, experimental set-ups and other methods contribute to the establishment of ratings.

Who will use the ratings? How are the ratings distributed - are they for internal use of a company, a school, an institution or are they distributed widely and - at least in principle - open to the public?

Functions of Rating

Let us first start with a list of traditional forms of rating. We will classify ratings in the categories

- Assessment of performance of individuals
- Reduction of uncertainty in making decisions
- Assessment of output and achievement

One may also classify differently:

- Assessment of past actions and performance, or existing products and artifacts to serve as a model
- Assessment of alternatives (specific plans) for action
- Assessment of possible future developments, visions or states of the world
  (independent of the path to this futures),

It is obvious that these functions are not clearly separated. Many rating mechanisms have several purposes and different users of a rating will see different functions as important. However, to achieve some order in this area it could be useful to use this preliminary classification. We leave out the evaluation by the marketplace: prices are a form of rating. One the one hand they are indicators of the usefulness but they are also influenced by scarcity and the effort to produce them (expressed in capital and labour costs). However, it is self-evident that many of the rating schemes listed below are somehow - directly or indirectly - related to economic performance. We have to be aware that the different forms of rating have quite different social and political influence: ratings by big credit institutions may determine the future fate of companies and their employees. They may even topple governments in case a nation is loosing creditworthiness (Martin, Schumann 1996, p.99).

We will study those features in turn and will make some remarks on their implications for Computer Mediated Communications (CMC).

Assessment of performance of individuals

Grading in school and higher education

We will start with this as most of us have experienced this kind of rating as the first formal rating scheme to become acquainted with. It is perceived as a grading of the knowledge and the ability of the pupil or student. Empirical work (Kalthoff 1996, Stiggins and Conkin 1992, Brookhart 1993) shows the complex setting of grading practices. Teachers adapt their grading to the average performance of the class. They orient themselves on the best pupils and restructure their assessment according to averages (arithmetic means, number of pupils failing a given task). Teachers perceive the results of examinations as a test of their own performance. Have they been successful in explaining concepts and teaching skills? This becomes more explicit in objective forms of examinations, which is most often the case in final exams. Those are supervised by external experts (in most cases teachers themselve) and in some countries the tasks are determined externally. Comparison of grades by the instructor with those by independent (outside) experts provides clues for the ability and reliability of the instructor.

We have to learn from this that the responsibility of those who have to rate is dependent on their role in the production - evaluation/rating - consequences cycle.

Personnel evaluation

Similar reasoning is true for personnel evaluation. Assessment of personnel to be hired and on the job can be seen often as much as an evaluation of the evaluator as of the person to be evaluated. The literature on personnel management provides a number of different models which are implemented in varying types of industries. Performance rating on the job is another name of the same task. As an example ma serve the publication of the American Management Assosiation (Grote 1996)

Psychological testing

Well known and also used in personnel evaluation are the different versions of intelligence tests. Introduced originally to measure the intelligence of psychically handicapped it had widespread use in the seventies and eighties. Although the enthusiasm has somewhat diminished it is still widely used.

Following this psychologists and to some extent sociologist have devised a large number of experiments to measure psychologival parameters of persons, and to rank them on dscalles expressing various properties.

Evaluation of crimes

The evaluation of crimes is a typical multistage process: First one body, the legislator formulates guidelines for the lengths of containment in jail or the fines for conceivable crimes or offences - or an existing tradition of cases, textbooks and practices provides this framework. Then in a given case another body of persons - the judge, a jury - has to determine the specific penalty which takes into account the special circumstances of the offence. Even the mitigating and aggravating circumstances have legal codifications, precedences and a body of literature (e.g. Pallin 1982).

The legal system codifies to some extent the moral convictions of the society and provide means to enforce them. The purposes of the punishment are various: Prevention of future crimes and compensation or punishment for misbehaviour are the major factors contributing to the justification of these most elaborate rating systems of our societies. The quoted text by Pallin describes these factors in more detail. A very specific detail in this form of rating is the elaborate justification which goes with every verdict. If voters had to express their decisions always - or at least more often - in terms of well established principles consensus may be more easily achievable.

Even if one does not (yet) institutionalise the legal system in the electronic environment, it gives an example of the rating of undesired events and results. The rating of content on the WWW and in the Internet as pornographic or politically censured may provide a first meeting point of the net technologies and the legal system. The

multinational nature of the networks will make progress in this area dependent on long-lasting international consultation and negotiation.
This does not hinder individual rating agencies to provide voluntary tools to discourage unwanted content.

Reduce uncertainty in making decisions

Product and service rating (consumer associacions, special interest and technical journals)

The increased diversification of goods and services makes it more and more difficult to choose products according to functionality and quality. Internationalisation of supply and regionally diversified marketing strategies make the orientation of the consumers, but also of professional users of investment goods and services ever more complicated. This has led to emergent consumer associations and professional consulting services.
These consumer associations, professional consulting companies, technical journals and special interest magazines publish regularly reviews and comparative assessments of consumer products and professional working tools. The outcome ranges from vague recommendations to tables listing many different features of the items compared,. The different features are often combined (most often by weighted sums) to achieve an overall assessment. This is done by (hopefully) independent institutions and journals.
The producers or providers of the respective goods or services use those results intensively if they are favourable. In regions where it is permitted they set up similar tables themselves - using only those features which are favourable to their product. The declared aim of independent rating institutions is the unbiased information of users and consumers. However, consumer associations report (Spitalsky 198) that the producers react on the results of the reviews and change product designs and features.
This area of direct influence on sales is therefore often put under some pressure from the side of the rated organisations. It shows that objective tests and assessments of this kind needs some proof of validity and trustworthiness. The trust in organisational arrangements or in individuals doing such comparisons is built up only slowly. In the end the institutions have to provide consistently useful ratings to be accepted by the public or the professionals concerned.

Credit rating of individuals, companies and political institutions

Banks and suppliers of goods and services have a need to assess the creditworthiness of customers with whom the do business. Suppliers use very often information from banks and financial institutions to cope with unknown companies. Specialised service companies collect and assemble information of this kind. The procedures range from elaborate assessment of the economic performance and standing of companies or countries to over the counter lending by using simple questionnaires with computer support to assess creditworthiness for small scale debts.

As an example we will show here rating definitions of the large international rating institutions which rate (inter alias) the banks themselves (a3eco 1996):

Rating Definitions

Moody's		Standard&Poor's		Thomson Bankwatch				IBCA		Nippon	
short term	long term	short term	long term	countries	firms	short term	long term	short term	long term	short term	long term
P-1	Aaa	A-1+	AAA	I.	A	TBW-1	AAA	A1+	AAA	a-1+	AAA
(Prime 1)	Aa	A-1	AA	II.	A/B	TBW-2	AA	A1	AA-	a-1	AA
P-2	A	A-2	AA	III.	B	TBW-3	A	A2	A	a-2	A
P-3	Baa	A-3	BBB	IV.	B/C	TBW-4	BBB	A3	BBB	a-3	BBB
NP	B	B	BB	V.	C		B	B	BB	b	BB
(not prime)	B	C	B		C/D		CCC	C	B	c	B
	CAA	D	CCC		D		CC	D	CCC	d	CCC
	CA		CC		D/E		D				CC
	C		C,CI,D								C,D

Thomson Bankwatch has separate rating schemes for countries (governments) and companies (firms). Similar but often simpler rating schemes apply to credit rating for companies in the national environment. Credit rating are directly related to interest rates charged. The rating of countries by Moody's may increase interest beetween Aaa (Triple A) and B by 3.8 percent (Martin, Schumann 1994, p.98).

The process of credit rating by banks is quite elaborate. All kinds of data about the individual or the company are assembled and complex calculations based on statistical reasoning and the latest developments in mathematical decision making under uncertainty are used. Banks themselves use computer networks to refer complicated and risky decisions to central authorisation units. In this case the central unit and the manager in the outpost knowing the local circumstances can have the same information on the screen and often decide together on the rating.

Evaluation of project proposals

Together with the evaluation of projects this is the area of the classical evaluation. Guba and Lincoln (1993) describe four different steps in the development of evaluations (their background is in educational projects whereupon the theoretical background of evaluation draws heavily):
Measurement (psychological - educational tests): Evaluator as technician
Description: characteristics of achievement and non achievement. Strength and weaknesses in the attainment of specified objectives.
Judgement: drawing conclusions about evaluands success, effectiveness or utility
Constructive negotiation (see below in the methods section)
For the time we may state that for the phase of project formulation and proposal evaluation communication is of high importance. Not all discussions can be substituted by CMC but it helps in the follow up of personal discussions and presentations. Reformulating in the proposal situation is well suited for negotiations with - and involvement of - special interest groups.
Clear cut ratings are possible but have to take into account the sensitive issues.

Evaluation of medical interventions

The assessment of medical treatments is a sensitive issue. The traditional double blind experiments try to avoid some common pitfalls of introducing bias by the experimenter. While a certain freedom in interpretation remains the procedures are mainly based on measurement and experiment, which is the first method of evaluation in well defined circumstances.

Distribution of resources according to urgency of different options

Performance rating of companies in the management boards and in the stock exchange

As we declared that we will not tackle the area of prices and economic competition there remain only the non-economic areas to assess companies: like rating companies according to their social impact or the ecological consequences of their actions. Rating of this kind is for instance performed by the US organisation Eco-Rating International which assesses companies and projects under an ecological perspective for potential ecologically aware investors. A special emphasis lies on Eco-Agro Rating.

Agenda setting in the political arena

Agenda setting is a political process full of pressure groups, social partners, subcutaneous influences of different kind. It is the area of what we call a negotiation environment for programme assessment and evaluation. New electronic media can contribute in the public discussion aspect of this agenda setting.
An example how ratings can be used in this area is provided by VoteSmart, an institution which tries to deliver impartial information for USA voters. One of its approches is to show ratings of the members of the US Congress by different special interest groups. The idea (independent of the Web or the Internet) is to count the number of cases the member voted with the interests of a given interest groups. This kind of performance evaluation is an established practice by those groups.

VoteSmart collects evaluations of this kind by many different interest groups and describes the results thus:

'These evaluations are in percentage form. They represent the percentage of time that the incumbent voted with that organisation's preferred positions on a number of votes that they identified as key in their issue area. Remember, by definition, these ratings by special interest groups are biased. They do not represent a non-partisan stance. In addition, some groups select votes that tend to favour members of one political party over another, rather than selecting votes based solely on issue concerns. However, they can be invaluable in showing where an incumbent has stood on a series of votes over a year's time, especially when ratings by groups on all sides of an issue are compared. Descriptions of the organisations offering performance evaluations are available'.

The effort shows how even biased ratings may be used to get a comparatively clear overall picture of an area filled with subjective and interest loaded judgements. The message to us could be: Always watch who is rating. Ratings themselve give clues about the rating unit (be it a person or an institution).

Assessment of output and achievement

Measurement of performance

In some subjects in sport and in competitive games easily definable procedures help to measure performance and rate achievement. Not too many of those achievements are related to activities in the digital environment. It is, however, no conceptual problem to show results on a new medium and to consider new forms of competition acting directly on the networks.

Evaluation of projects

This is the evaluation of a project during and after its implementation. Comparative evaluation of completed projects in the political arena are not popular but can give valuable insights in the planning of further activities.

Literary criticism - selection by publishers - peer review

Judgement of texts by publishers readers and by the literary critiques determines the success of texts. That is true for both the success in the market place and also the recognition in the more esoteric circles of elitist literature. A somewhat more formalised and theoretically more impartial solution is the peer review process which tries to guarantee a fair selection process. It is still the best process we have although it is sometimes distorted by influential individuals or by unscrupulous groups of those. A recent article in the Scientific American it is shown how difficult it is for scientists from third world countries to publish in journals quoted in the Science Citation Index, and how difficult it is for Third world journals to be accepted in the Index (This Index works obviously contrary to the former Vatican Index by exclusion - not by inclusion).

Honours, medals and prices (literature, film, arts, sciences)

Ratings are constituted fairly frequently in the different forms of art and in sciences by awarding prices. These range from fairly local events with limited appeal to outsiders to events with world-wide reputation like the Oscar, the award of the American Film Academy, and the Nobel price.
These two events show also the main sources which establish those events and keep them alive: institutions which want to promote a cause, e.g. an industry (in our case the film industry), a public interest (if the sponsor is a national or local government) or individuals which have a special interest in some area - and Nobel is by far not alone in doing it. The Nobel price became famous because of the money which came with it. This does not mean that it would not remain famous if the money would suddenly not be available any more. Some of them are voted for by large groups of people and represent therefore a certain consensus within the group.
Not always is the price necessarily only directed to its professed purpose. Often more or less veiled purposes - of political, manipulative or even tax evasive character - may be present also.
Prestigious prices definitely enhance reputation and are directly or indirectly also sources of economic benefits.

Achieving specific goals: social, technical, environmental

Professional societies recognise the achievement of some of those goals by award or prices. Other achievements are honoured by the government in giving tax benefits, grants or contracts.
Honours and prices are one form. Some organisations or publications, however, list institiutions and persons also indepentend of those. They praise those doing active work for a cause, or behaving in the right way, while other lists single out the worst offenders of the expected behaviour.

Peace and human understanding

We take this as an example of widely recognised but often difficult to define contributions. Receiving a peace price does not prevent recipients to go to war later on. It may also be the token for abstaining from further horrors. But do not misunderstand me: This is an achievement!
Increasing awareness of refugees, the hungry of the world, the dangers of war and the mechanisms leading to those is worthwhile pursuing even if it sometimes goes awry.

On the opposite end there are also medals and appraisals for acts of war, the war heroes as outstanding examples. Rating as an abstract concept is neutral. As individuals or as institutions we have to take our stand.

"Over all" achievements. Place in history

The long term rating of achievement is the recognition in the text books of the area concerned, and the inclusion in general books of history. While even here some mechanisms to remember and honour your own kin are present (those winning the wars write history) a certain detachment allows a more sober view. The sheer need to concentrate on the essential contributes to a fairer rating. History writing may surface contributions which were not recognised as forerunners of important developments during the life of the originator(s). But history rarely rectifies misjudgement during the lifetime of able contributors and in a very precise sense it comes to late for those anyhow.

Methods of Rating

Methods of rating are influenced by a number of factors:
Rational measurement and assessment of performance. Using experiments and related observations, measurements, mathematical and statistical models to simulate or replicate behaviour. Discussing arguments in the context of the knowledge to which the rated object or person contributes.
Tactical and strategical considerations are included into the rating to achieve a political, personal, or group oriented objective. Sometimes the real purpose is somewhat hidden and the rating assumes a manipulative character.

Measurement and experiment

We have already spoken about measurement in sports. Testing products is another area where many features or ingredients of a rating may be determined by experimental set-ups. These set-ups may also include users, giving their subjective opinion on some quality. How far, and under what circumstances this is done, determines the answer whether the whole procedure is still a measurement. The problem is then more how to aggregate the different results in an overall rating for a specific purpose. As long as the individual (partial) test results are made accessible, alternative ratings for special purposes can be derived. An appropriate digital environment is well suited to support this re-evaluation.

Specialists. Opinion leaders

Many, if nor most formal ratings are done by experts and specialists. They try to establish formal procedures to reach an objective result. These may be experimental set-ups in which consumer products or other items may be measured.

Thus, the International Organisation of Consumer Unions, IOCU, publishes guidelines for testing. The European Testing Group (ETG) organises co-operative testing of products by the different national testing organisations. Additionally to the staff of the institutions experts are mustered for the special area of the product under consideration.
The properties of the product which are considered are functional quality, durability, safety and security, ease of handling and price. Especially, properties offered in advertisements, legal requirements, standard, environmental compatibility, service, warranty, availability of spare parts packing installation and manuals have to be taken into account.
Experiments show that independent evaluation of products by consumers highly correlate with the test results. A certain problem is the overall assessment. This global rating is assuming an average consumer and may not fit to the individual need of the specific consumer. This discrepancy is accepted to achieve high visibility and to reach consumers with widely varying education.

If the aim is to predict, based on queries to a sample of the whole population, the choices which would have been made if every single member of a large population had been asked, then statistics requires stringent ways of selecting the sample, using random sampling, stratified sampling, etc.

Jury

Prices, medals and honours (and legal sentences) are often given by juries. In most of the cases the jury is a group of experts in a given field.

Consider the Nobel price awarding procedure:

Each year the respective committees send individual invitations to thousands of scientists, members of academies and university professors in numerous countries, asking them to nominate candidates for the Nobel Prizes for the coming year. Those who are competent to submit nominations are chosen in such a way that as many countries and universities as possible will be represented. These prize nominations must reach the respective committees before February 1 of the year for which the nomination is being made.

'The nominations received by each committee are then investigated with the help of specially appointed experts. When the committees have made their selection among the nominated candidates and have presented their recommendations to the prize-awarding institutions, a vote is taken for the final choice of laureates. Prize decisions are announced immediately after the vote in October each year. Competence to nominate candidates for the Nobel Prizes varies somewhat among the prize-awarding institutions...'

Juries somtimes work in very informal ways but some have also highly structured procedures. It may be conjectured that the selection of the jury is more important than the procedure, but in some cases well defined, published procedures are important for the acceptance of the rating by the users.

Peer groups

The paradigm of assessment in science is the peer review. What is to be published in influential journals is determined by anonymized scientific peers. The articles are sent by the editors of the journals to peer scientists competent to judge the content. These do agree or not to a publication. They often provide further guidance and helpful comments on the paper and make publication dependent on the meeting of certain conditions.

Due to the widespread use of Internet publishing in some sciences (50% of all physics papers are pre-published on the e-print-Archive in Los Alamos, New Mexico) the discussion on electronic alternatives to the peer reviews has advanced most widely. A whole OECD conference in June 1996 was dedicated to "The Global Research Village" (The Economist June 22 1996).

One solution would be to append comments to every paper for subsequent readers to view. Another to put stars on the paper, like in the Michelin Guide (one of the consumer product rating schemes). The chief worry is "that high quality work will be drowned in a flood of dubious data and poor prose." The peer review also lets a lot of this through.

Normally the author of the paper gives up his copyright to the scientific paper. Publishing on the Internet means keeping the copyright but giving it away for free. Should the government interfere? Similar to what the Danish minister of research and information technology proposes: We will provide high speed communication links if the scientists publish their results on the Internet!

Clearly a final model to replace peer review has not been found. Web4Groups i a place to experiment.

Social discussion, bargaining processes and negotiations of those concerned

Evaluation of proposals or of ongoing or finished projects are often performed by a 'political' process of negotiation of the parties concerned. The proponents of 4th Generation Evaluation (Guba and Lincoln 1989) argue that evaluation in such circumstances should prefer such a way of 'constructing' a joint view of the envisaged changes and of the effects on groups with differing interests. Such a process will often not be satisfied with ranking a limited number of proposals but will be actively involved in developing proposals.

Opinion polls

Opinion polls are often conducted to find out the preferences of the public or a specific target group. Normally a limited set of options are proposed and preferences are looked for.

The main problem on the Internet is so far representativity. Only target groups with strong commitment to the net will be appropriate for this method.

Automatic Procedures

In electronic environments many procedures may be automated: Users or visitors of a Web site or another entry point to a computer may be observed. Inferences about their preferences can be made automatically by analysing their path through the computer, the Web-site or the Web at large. This ratings can then be combined in an adequate way and be presented to the user himself (may be for correction or to get a more explicit rating from him) to the managers of the site or (at least in aggregate form) to the other visitors of the site ("Our most successful pages are". "This URL lured 20.000 repeating visitors" or more complicated inferences). There is some concern about the transparent customer about whom the shopping centre knows more than he himself. Should one encourage this? Should one give back some of the information to the user/customer who provided it in the first place? Where are the limits?

Stratified or general rating

Rating may even be done by individuals. However, in many cases the rating shall represent a totality of a certain sort. In this case one has the choice that everybody can contribute to the rating process of all items or issues. There is also the other possibility that the group elects representatives to do the actual assessment for them (zhis can also be a single individual). Alternatively a random sample from the totality can do this rating.

Stratified samples guided opinion formation

A variant of the random sample is the approach to assemble a random sample of the totality, selected by main characteristics of the population (stratified sample), and to perform with them a guided opinion formation. Assuming that the average person selected will not be too competent in the area, experts will present the main positions and issues in the area for which a rating may be due. After the presentation and guidance the group is left alone to come to a common conclusion in a negotiation phase.

Contributions of voluntary donators

Simple rating tools on the WWW give the opportunity for every visitor to leave his preferences on a limited number of topics or items. Those are the assembled in overall ratings or statistical distributions of results.
This resembles street activists who collect opinions on debated issues by asking passers by to answer questions or fill in forms. It has some justification if the visited arena is rather specialised and its visitors are reasonably well behaved. It may also be suited to make visible a
public outcry or the complete negligence of an item.
If the aim is to predict, based on queries to a sample of the whole population, the choices which would have been made if every single member of a large population had been asked, then statistics requires stringent ways of selecting the sample, using random sampling, stratified sampling, etc.
However, in many everyday processes, there is no such goal to predict the opinion of a larger population. Quite the opposite, it is often valuable to have rates set by individuals who have time to prepare a good evaluation before setting their ratings In this case it may be desirable to identify the persons giving the ratings and make them recognisable.

The fact that statistical predictions requires stringent sampling methods does not imply that all kinds of ratings done by only a selected number of people to be regarded as non- acceptable. In any case one will gather with these methods ideas about what may be controversial or where there is little dissent.

Commercial success in the market place

For completeness we mention this most widespread and successful rating scheme. It will not be dealt here explicitly but it enters in other ways implicitly.

Distribution and Dissemination of Ratings

Confidentiality - distribution policies

The distribution or dissemination of the results of a rating process varies widely. Examination grades, personnel ratings, credit ratings are not usually distributed widely. Even if they are announced the circulation is fairly limited. Privacy and Data Protection laws even forbid the publication of many of those informations.

Research results, sport events, tests by consumer unions and prices and honours are for publication and the amount of publicity depends more on the influence of the rating unit on media than on demands for confidentiality.

Many ratings inside government institutions are confidential, at least as long as the final decision is not announced.

Every distribution mechanism will have to take into account the different needs of diverse rating institutions. Some demand for confidentiality, especially in the phase of collecting contributions to the rating process will be necessary for many organisations. On the Internet the security features which are just stabilising will be necessary to provide this data protection.

Who receives ratings

Different phases of the rating process requires changing needs for the protection of information. Some prices are awarded very openly: discussions on the merits of the candidates may even be transmitted over TV. Other large organisations (think of the Oscar) keep intermediate results fairly concealed. And many ratings have to be accessible only to a few select.

In a distance education environment one may ask that grades are only visible to the examiner, the school administration and the student. In many of those circumstances the rating of the rater may depend on his successful ratings.

Who uses ratings

The different functions determine the target groups of the ratings.

Rating students, personnel, and crimes the target groups are obvious. Only large crimes are published widely although the legal proceedings are normally public. In Europe at least it is not usual to put criminals on display as a part of the punishment. But both the police system and potential employers are interested in some information.

All ratings directed to decision makers are evidently used to help making those decisions.

The most diffuse motivation exist for the assessment of output and achievement. Part of the raison d'être lies in giving and enhancing reputation for those whose achievements seem worthwhile to one group or another. That implies that the ideas the results or only the biography of the rated person, organisation or the project become better known. Here not decisions are in the forefront but orientation in a longer time-frame. Those providing these ratings hope to advance a cause, to share enjoyment about achievements and to contribute to the structuring of our life. Decisions are influenced, but those giving the award have no specific idea about the individual decisions made.

Projections

The comparison of ratings by several or many subjects may be used to find common interests between persons. Having identified a common interest (e.g. in music, in science, in literature) it is possible to make proposals for interesting items one of them has seen and rated.

How long are ratings significant

Different forms of rating have different life time. Normally ratings are superseded by more recent ratings of the same kind. But even if their validity remains intact their impact may change over time. Consumer product ratings would often be valid several years. Analyses of the impact of especially good ratings for products are felt by the sales departments of the respective companies between 3 and 7 months (Spitalsky 198x).

Support of Rating through CMC

There are some differences between face to face meetings with (synchronous types of) voting. Computer support allows the handling of much larger numbers of participants. More complex sequences of questions may be designed to give answers to several aspects of the problem under discussion. Several and more elaborate aggregation rules may be used and discussed. Computer support allows also continuous voting, where every participant may see the result of the votes up to now and may change his vote accordingly. This allows decision procedures akin to Delphi studies in which experts adapt their estimates of future events and of ratings of options under the influence of arguments of other panelists. Often but not always a better consensus may be achieved.

Role distribution (Who is rating?)

Access control present in most CMC tool may be used to restrict access to the rating process. Outcome of any rating process is crucially dependent on who is doing or contributing to the rating.

Interest groups have usually a formulated purpose and rate according to their interest. Normally they have also an internal structure which can define appropriate roles and procedures.

If we want to implement a rating mechanism which every participant in CMC can start, we will have to define a minimum set of those roles and a minimum procedure.

If we are more ambitious we can devise a whole class of procedures, each with its own role set. In any case we can assume that anybody setting up a rating scheme will do the necessary preparation. Either using the predefined (minimum?) roles and procedures or customising the server for his needs. Computer help could especially be provided in the aggregation process of individual contributions.

Agenda Setting (What is rated?)

Professional rating institutions have well defined tasks and established ways to select items to rate. The ease of communication could make it simpler to handle larger volumes of data. The countervailing force is the need or hope for a certain quality of the rating. Otherwise the acceptance and the use of a rating scheme will be missing. The alternative could be the manipulation of results by the interested parties.

One has to be aware that already the decision to rate something is an important one. Even a film getting the predicate of the worst film ever produced gets some additional visibility and some fame after all. Even more important is agenda setting in the media. Creating awareness is a separate activity which may be supported by nearly any rating procedure.

Monitoring (What are the others doing?)

For many purposes it is enough to follow the activities of others. If I have identified a person or an institution who does excellent work in my area, it is advisable to follow the activities of the person or institution to the extent which is manageable for me. Does this need active involvement of the expert? Or should we look only to changes of his Web (or Web4Group) site?

Filtering and Prediction (What are my interests?)

A basic problem on the Internet is the mass of information. To find items of high quality fitting exactly my interest and my pressing needs for searched for items becomes more difficult in the present fast growing Internet. We described already the possibility to use joint interests to make predictions about items which could interest me.

Simpler methods just filter out unwanted information or direct me to the areas which I visited before.

The theoretical problem is to classify the information available not by the needs of an average user but to the needs I have. If I consider my interests and knowledge represented by a classification, I am looking for a classification of the material on the Internet which comes near to the classification I am acquainted with. Naturally I do not need the information which I already have, but some which is near enough to be useful but still complementary. To do this automatically is a topic for research in statistics and artificial intelligence but practical results are still far away.

On the other hand it remains likely that the most important way to find out about new relevant developments is the personal contact with other persons with similar - but not too similar - ideas, interests and background. Additional labels or ratings by special interest groups could help to narrow down the field to search for specific information. The traditional keepers of information have been the librarians. They have awealth of experience in classifying information for users. They were struggling with the traditional tools, catalogues and card indices. But they have developed many ideas going beyond. One has to transfer several of the virtues of librarians into the cybersphere. The chapter on existing rating tools will describe first applications.

References

This paper is based on the relevant content of our book
Alton-Scheidl, Roland; Rupert Schmutzer; Peter Paul Sint; Gernot Tscherteu:
Voting., Rating, Annotation. Austrian Computer Society, Vienna 1997
Chapter A.2.2. Rating in everyday life

a3eco 1996; Rating an der grossen Glocke, a3eco 1996, p8

Bierman, Todd; Nathaniel Wice: The Guerrilla Guide to Credit Repair 1994: How to Find Out What's Wrong With Your Credit Rating-And How to Fix It. New York: St Martins Press

Bortz, Jürgen & Döring, Nicola 1995 (2. überarb. Aufl.): Forschungsmethoden und Evaluation; Berlin Heidelberg New York

Brookhart, S.M. 1993: Teachers Grading Practices: Meaning and Values. Journal of Educational Measurement. 30: 123-142

Buzzard, Karen 1992: Electronic Media Ratings : Turning Audiences into Dollars and Sense, Focal Press, Newton, MA, USA

Eco-rating: http://www.eco-rating.com/

Grote, Dick 1996: The Complete Guide to Performance Appraisal. New York: Amacom (American Management Association)

Guba, Egon G; Yvonna S. Lincoln 1989.: Fourth Generation Evaluation.- 1.print. - Newbury Park, Calif. [u.a.] : Sage Publ

Ingerling, Richard 1980: Das Credit-Scoring System im Konsumentenkreditgeschäft: Konzeption und Wirkung eine Rationalisierungsmittels in der Kreditwürdigkeitsprüfung. (Gundlagen und Praxis des Bank und Börsenwesens 12). Berlin: Schmidt 1980
(Credit-Scoring System in Consumer Credit Business. In German)

Kalthoff, Herbert 1996: Das Zensurenpanoptikum. Eine ethnographische Studie zur schulischen Bewertungspraxis. Zeitschr.f.Soziologie, 25 (2), April 1996, S 106-124 (Figure Cabinet of Grades, an ethnographic study of evaluation practices in schools. In German)

Martin, Hans-Peter; Harald Schumann 1996: Die Globalisierungsfalle, Rowohlt, Hamburg

OECD 1994: Evaluation and the Decision Making Process in Higher Education: French, German and Spanish Experiences. OECD Documents, Paris 1994

Pallin, Franz 1982: Die Strafzumessung in rechtlicher Sicht. Wien, Manz 1992 (Determining the Size of Punishment in Legal Perspective. In German)

Spitalsky, Hann es 1995: Der vergleichende Warentest als konsumentenpolitisches Instrument
131-140 (Comparative Testing of Products as Instrument of Consumers Policies. In German, available at Österreichischer Verein für Konsumentenberatung, Wien)

Stiggins, R.J.; N.F.Conkin 1992: In teachers hands: investigating the practices of classroom assessment. Albany: SUNY Press

The Economist 1996: Re-engineering peer review. The Internet. June 22nd 1996, p98-99

Application of a Generic Voting Tool for Rating Purposes

András Micsik
micsik@sztaki.hu

Department of Distributed Systems
MTA SZTAKI, Computer and Automation Research Institute, Hungarian Academy of Sciences
H-1111 Budapest XI. Lágymányosi u. 11. Hungary

Abstract: A highly customizable voting subsystem has been implemented as part of the Web4Groups EU supported project (Telematics Application Development Projects, Fourth Framework Program). The target of the Web4Groups project is to develop a distributed non-simultaneous group communication system with multiple access possibilities (WWW, mail, fax, etc.) and incorporating advanced groupware functionalities such as voting, rating and annotation. This paper describes the voting facility of the Web4Groups system, and investigates the application of this voting tool for rating purposes.

Keywords: CSCW, groupware, group communication service, World Wide Web, voting, rating, PICS, Web4Groups

1. Introduction

Rating can help people in multiple ways to navigate on the Internet more effectively and safely. The need for practical rating facilities over the Internet is shown by the emerge of the Platform for Internet Content Selection (PICS, this effort is guided by the World Wide Web Consortium). PICS defines the content and the communication of ratings among Internet hosts. In PICS terms a label bureau serves the ratings provided by one or more rating services [5]. Rating services give their ratings according to rating systems. A rating system defines the syntax and semantics of the possible ratings [4]. A separate WWW page may define the meaning of the rating system for humans.

Presently the distribution and service of ratings seems to be solved, but there is a lack in tools and unified environments for the collection and calculation of ratings. The Web4Groups system with its wide group collaboration features could help in the operation of rating services. To achieve this, relatively small extensions are needed to the existing Web4Groups software. This paper briefly describes the Web4Groups system (Section 2), and its voting facility (Section 3), then investigates the possibility of extension for rating purposes (Section 4), and finally gives some scenarios for the use of the extended system (Section 5).

2. The Web4Groups conferencing system

Web4Groups is a distributed system that has a notion of users, documents or messages and activities. Documents can be stored under activities. Users can browse the mesh of activities, viewing and adding messages inside activities. Activities can be of different types with different behavior or additional functionalities. Currently the most important supported activity types are:

- forums for public discussions
- workspaces for limited access to a given set of users
- votes
- annotations for WWW pages

There is also an activity type for joint editing of compound documents under preparation. The basic group collaboration services of Web4Groups consist:

- group membership administration
- user authentication and authorization
- personal workspace management
- multimedia E-mail support
- support for multilinguality

Another strong side of the Web4Groups system is the multiple ways of accessing its information. Currently besides the WWW user interface there are also user interfaces under implementation for telnet, telephone and fax connections.

The architecture of the system is based on a special database called KOM that stores objects of the system (messages and activities) called "boards". Boards can be connected with typed links to each other enabling a highly flexible structure for the groupware functionality. The database has been implemented at SICS in Sweden in C++ language.

The different user interfaces (WWW, telnet, phone, etc.) are separate software entities communicating with the KOM database via a TCP/IP based protocol. As the system is distributed, the KOM databases have another protocol for their inter-database communication. The WWW user interface is implemented in Java by Kapsch. For more information about the Web4Groups features and functions please refer to the published information about the project and the system. [6,7]

3. The voting subsystem of Web4Groups

Voting is integrated into the general conferencing features of Web4Groups. A voting is presented as a set of Web4Groups forums and messages. Special actions in a voting are shown as buttons when a user browses the voting. This way, user registration, access permission, message threads, multilinguality and distributed behavior are inherited from the Web4Groups system. The voting subsystem is implemented as a plug-in module for the Web4Groups system. It has been developed in Java language by SZTAKI in Hungary.

In the design phase of this subsystem SOCOEC (Austria) prepared an in-depth study on the use, mechanics and social aspects of voting and rating processes in everyday life [8]. This implied the idea of a generally applicable voting tool [1,3,9]. This tool controls voting processes according its configuration given by the vote organizer. The configuration includes:

- definition of user groups
- definition of the questionnaire (vote form)
- definition of the voting process

Members of user groups are defined in such a way that every member can have a selected language for communication with the system. Questionnaires can also be given in multiple languages, thus continuing and improving the multilinguality aspects of the Web4Groups system. A large variety of question types are supported including single or multiple selection, evaluation with given labels and ranking.

The control flow of the voting process is defined by a state-machine (called script). This state-machine can act on conditions such as a new vote's arrival or the value of a system maintained variable (e.g. 90% of the participants has voted). It can also perform actions at a given time (e.g. to stop the voting at midnight). Conditions fire the execution of command blocks. Commands cover all generic actions during the voting process in a simple manner which does not require programming skills, neither allows the abuse or corruption of the vote. There are commands for sending messages to groups of people, publishing the results, changing vote switches, etc. Switches provide general configuration of the vote, defining user authentication, anonymity, or a statement whether one may change his/her vote during the voting process or not. For the configuration and management of voting processes a WWW form-based interface is provided.

3.1 The course of a voting from the users viewpoint

At the time of creation, a new workspace is generated for the vote with special messages (e.g. description, log) and a separate subforum for discussion. The organizers of the vote are appointed, and they perform the configuration of the vote considering the remarks and discussion of participants. The configuration of the vote and the vote form is shown in the workspace.

When everything has been settled, the voting process may be started. Participants may vote either by filling the questionnaire via the WWW interface, or by sending their ballots in e-mail. In this phase normally no interaction is needed by the organizers, but in case of disorders they have a possibility to interrupt and fix the voting process. According to the script invitations, reminders, results are generated automatically. Finally the voting process is terminated and the vote is closed. This process can be followed in the log (which is a message readable to all participants). After closing the vote, the workspace turns into an archive, storing all important documents of the vote.

4. Extending the voting subsystem for rating purposes

In the first subsection the possibility of such extension is studied. Then further subsections describe one way of performing the extension.

4.1. Comparison of voting and rating

A rating process can be seen as the same voting process cloned for each object to be rated. For each rated object a separate voting is performed, but these votings share the configuration, i.e. the questionnaire, the group of allowed voters, etc. In this view a voting process can be extended into a rating process by adding a new dimension to it namely the group definition of rateable objects. To refine this view the significant differences between a rating process and a voting process are summarized here.

Differences in the definition of voting and rating processes

Rating introduces the task of associating an object with a rating. The first aspect of this is the definition of rateable objects. The group of rateable objects can be limited according various attributes, for example format, location, or topic.

On the other hand definition tasks inherited from voting are richer than it is expected for rating. Complex time schedule is rarely used in case of rating. The only essential control flow is that after the arrival of a new rating certain commands are executed. Among question types found in votings mostly single and multiple selections are used.

Differences in the result calculation

In case of rating the rated objects are ranked with respect to their rates. In case of voting the choices of a question are ranked according the votes. This calculation can be used to produce the rating of one object. A second level of computation has to be added to compare the ratings of objects and to get different rank orders of the object. Like there are several methods to calculate the result on the first level (the rating of one object), there are also several methods to calculate the rank orders of rated objects on the second level.

Distribution of the result has different methods

In a voting process the result is treated as a whole, while in a rating process the rates are queried either individually for each object, or in complex database-like queries. Rated objects are retrievable in various sorted orders or by query formulae. Rating results are usually distributed through a query interface and not as huge rank lists of objects. For such interfaces different query types, query engines and distribution formats (for example PICS) can be used.

User interface

Rated objects and rating processes are interconnected with several relationships. These relationships has to be visualized for the user in an easily comprehensible way. For an object the available ratings and rating services has to be shown. If the user is allowed to contribute his/her own rating to a rating system, this has to be offered in the user interface while viewing the object. On the other hand a rating service may provide various rank lists of rated objects in which case the rated objects are to be uniquely identified and easily retrievable from the user interface.

Common features

In spite of these differences the definition of user groups, questionnaires and control flows can be used in both rating and voting processes. Similarly the methods used in connection with questions (presentation, filling in, result calculation, etc.) are common. Going through the list of features implemented in our voting tool none of those - though rarely used - proves to be useless. As a conclusion it can be stated that a general rating facility can be specialized from a general voting facility by inheriting all features in the voting tool and providing additional mechanisms for ranking the rated objects and for the association of rated objects to rating services. On the user interface side this may include a rating viewer/composer for objects and a query interface for rating services.

4.2 Accessing and submitting ratings

After the above examination the additional features and user interface elements are elaborated in the context of the Web4Groups system.

While viewing Web4Groups boards (internal objects) the user may ask for a separate rating window, where all information concerning rating is shown. For each available rating service inside Web4Groups the window shows:

- general information about the rating service (full description is available following a link)
- the actual rating of the viewed document (if it is available)
- the user's own rating for the viewed document (if it is available)
- a link to the page where he/she can submit/change his/her own ratings (if the user has proper rights)

Furthermore if the viewed object is a forum or workspace, then the above information is presented for the contained messages as well. For each rating service in concern the list of rated messages together with their ratings are provided.

4.3 Presentation of ratings

Presentation has two main tasks: to show the rating of one object in a very informative way, and to produce different rank orders of the objects. Presentation of the rating of one object raises the question how to show the most information in the less space. The rating of an object has to be quickly recognizable, though it must not take much of the space available for showing the whole object. In a listing of objects it is even more critical to compress the rating information. The best solution can be to assign icons to rating labels (for example 3 stars means very good, a trashcan means very bad). This shorthand notation can be a single word as well. However these shorthands hide many available information: how many people have rated the object, what is the distribution of the rates. A miniature histogram or bar chart of ratings may provide more complete information.

Rank orders are objects listed according to some of the rating categories. Each rating category can define a separate rank order (for example quality or genuineness). The basic operation behind ranking is the comparison of two rates. Rating services with a given goal may evolve their special ranking algorithms to provide the best rank lists for their users. In a general rating tool it is desired to find a generally applicable ranking algorithm, or to offer various algorithms for ranking which needs further investigation.

A tabular format can be devised for the visualization of ratings for a group of objects. Each category has a column, and cells contain an appropriate chart or icon for the rating. Clicking on the header of the column ranks the objects according to that label or category.

4.4 Query interface

Another way of discovering the ratings is through a query interface. Each rating service is accessible for the users via this interface. Here the users can read the detailed description of the rating service, and search the ratings of that service. For example the top ten objects according to the first category can be retrieved. The interface gives back the title and the URL of objects as a link, so the user can download a selected object into the WWW browser.

4.5 PICS interface

This interface is designed for machine-machine communication, ensuring that any external software may access and use the rating information stored in the Web4Groups system. The PICS recommendation provides a common language for the communication of rating information. The most important supported methods are:

- get the definition of the rating system in PICS format
- get the rating for a URL in PICS format
- get the list of rated URLs

4.6 Setting up a new rating process

Setting up a new rating process will be supported with a wizard, which guides the user through the following steps:

Give a description about the rating service

A textual description about the purpose of the service, how it works and who operates it.

Define rateable objects

This is done by setting allowed and disallowed object types or URL prefixes.

Define raters

The user group definition methods of the voting subsystem are applied here.

Define questions

Questions and their evaluation are composed with the form editor of the voting subsystem.

In case of non-public rating services the access to the service can be restricted by allowing or disclosing Web4Groups users and external connections from given IP domains.

5. Possible usage scenarios of the new rating tool

5.1 Rating of documents inside the Web4groups system

Scenario 1: Papers submitted to a workshop are rated for acceptance.

All papers are uploaded into a Web4Groups workspace, where only the authors and reviewers can access them. A new rating is configured exclusively for the objects in the workspace, with the reviewers as allowed raters. The review form is defined for the rating. Every reviewer can read the papers, and fill in the review form for the paper. The summary of the reviews containing the average points gained and the list of the reviewers' comments are automatically generated and can be seen by both the reviewers and the authors. After the review the rating process is stopped, and the rates are frozen in the workspace.

Scenario 2: The most excellent contributions are searched on a Web4Groups server

Many times it is a relevant need to find the best quality pieces in a large pile of information. In this case a new rating can be set up for a whole server. The rating questionnaire can be very simple and pinpointed at measuring the quality of the object. Any user of the server can fill in the questionnaire and affect the overall rating of the object. In this scenario it is not enough to present the rating for each object, but it is also important that object can be sorted or searched according to their quality rating.

5.2 Rating of WWW pages

Scenario: Implementing a PICS-based label bureau

A Web4Groups server extended with voting/rating can host several rating services for the Internet. Each rating service creates a working area containing public and private documents about its operation, and a separate rating process providing the operational functionality. The rating system can be defined as the fill-in form of the rating process. The definition of the rating system is automatically generated in PICS format. The raters get an account on the Web4Groups server. While they are connected to the server, they can enter their ratings for any WWW page. The results of their ratings can be asked from the server by anybody. The results are sent in PICS format to the user of the rating service.

Bibliography

[1] M. Biro and L. Kovacs, A. Micsik and T. Remzso, Reference Model for World Wide Web Voting and Rating Services, 7th Mini Euro Conference, Bruges, March 24-27, 1997

[2] Annex A: Draft user requirements for computer-supported voting and polling as a group communication task, (ISO/IEC JTC 1/SC 18/WG 4 N 1851 annex to N 1847)

[3] G. Kiss and L. Kovacs and A. Micsik, Voting and Rating within Web4Groups, First European Conference on Voting, Rating, Annotation, Vienna, 21-22 April 1997, (URL: http://www.web4groups.at/w4g/conf97)

[4] PICS Label Distribution Label Syntax and Communication Protocols, Version 1.1, W3C Recommendation, 31 October 1996

[5] Rating Services and Rating Systems (and Their Machine Readable Descriptions), Version 1.1, W3C Recommendation, 31 October 1996

[6] The Web4Groups Project (URL: http://www.web4groups.at/)

[7] Web4Groups base system manual, Deliverable D 3.3 of the Web4Groups project

[8] R. Schmutzer and P. Sint and R. Alton-Scheidl and G. Tscherteu, Voting and Rating - Perspectives for Information Collection, Decision Making and Collaborative Rating Using Web4Groups, Deliverable D 4.5 of the Web4Groups project

[9] L. Kovacs, A. Micsik, The Design of Voting and Rating Services within Web4Groups, CSCW in Design, Bangkok, 26-28 November 1997

Social Filtering and Social Reality

Christopher Lueg
AI-Lab, Department of Computer Science
University of Zurich
Winterthurerstrasse 190, CH-8057 Zurich
Email lueg@ifi.unizh.ch Fax +41-1-63 56809
URL http://www.ifi.unizh.ch/~lueg/

Abstract

In this paper, we argue that most current social information filtering approaches may benefit from more seriously taking into account the peculiarities of human cognition and human social behavior since current approaches only consider de-contextualized ratings. Social filtering systems exploit ratings provided by users in order to compute recommendations for other users. Typically, these ratings are detached from the situation and the social embedding in which they have been provided. Recent research on human cognition and behavior suggests that actions should not be viewed in isolation from the situation in which they occur (thus, the term "situated actions"). Accounting for the situation and the social embedding requires support for exploiting the situation rather than abstracting away the situation. In respect to exploiting the social embedding of ratings, we discuss the need for two related basic research directions. First, a self-organizing network of users trusting each other should be explored as a basis for true socially embedded filtering. Second, the suitability of collaborative filtering techniques as a tool for maintaining the focus of Usenet discussion groups by exposing spam and other clear off-topic postings should be investigated.

Introduction

Social filtering systems, also referred to as collaborative filtering systems (Goldberg *et al.* 1992; Resnick *et al.* 1994; Konstan *et al.* 1997) or recommender systems (Resnick & Varian 1997), aim at automating the "word of mouth" (Shardanand & Maes 1995). Relying on recommendations given by others usually happens in situations with either too much or too few information available. Prime examples for successfully implemented social filtering processes are people reading newspapers since these people trust in the decisions of the editors to include the most interesting and important articles. Recommendations for movies, compact disks, books, and events given by editors of journals to help their customers or recommendations given by friends to help friends are other common examples for relying on the judgments of others in unclear information situations.

Examples for the collaborative filtering approach, on the one hand, are systems filtering Usenet articles (e.g., Brewer & Johnson 1994; Resnick *et al.* 1994; Konstan *et al.* 1997; Terveen *et al.* 1997b). In the presence of a large amount of low quality items on the net, also called electronic junk (Denning 1982), the idea is that consumers help each other to distinguish between high quality and low quality items by providing ratings for items they have investigated. These ratings are collected and can then be used by others to focus on those items collectively rated best (or at least rated acceptable). Recommender systems, on the other hand, have been implemented in various domains, such as recommending webpages, music, or movies (e.g., Shardanand & Maes 1995; Terveen *et al.* 1997a). Despite less stressing the necessity of personal relations between the recommenders, the technique is basically the same as in the collaborative filtering approach.

Most social filtering approaches share some implicit assumptions that are explicated in the following. It seems as if the independence of ratings from both the topics and the representations of the objects being rated turns out to be the main lever applied by social filtering systems. Contrary to content-based filtering systems, social filtering systems are able to handle both virtual objects, such as Usenet articles or webpages, and real-world objects, such as movies or music, that are usually inaccessible to computers. In order to deal with ratings (a prerequisite for computing a recommendation) it is not necessary to analyze the corresponding objects as in content-based approaches. Also, the social embedding of recommendations can be abstracted away.

We proceed as follows: First, we briefly summerize why cognitive processes, such as rating, and socially embedded processes, such as recommending, cannot be replaced by "technical" processes without loosing certain peculiarities. Next, we introduce situatedness as a concept that appropriately accounts for the peculiarities of human cognition and briefly discuss "situated information filtering". `spynews`, a newsreader that supports situated information filtering, supports situated actions by avoiding to abstract away the context in which the user's "interest" occurs. This work has been focusing on individuals and their particular situation only. Social filtering as a community-based approach seems to be a promising complement to our individual-based filtering approach. Finally, in respect to appropriately accounting for the social embedding of ratings and recommendations, we discuss the need for two related basic research directions in social information filtering. First, a self-organizing network of users trusting each other may serve as a basis for true socially embedded filtering. Second, collaborative filtering techniques may be suitable for maintaining the focus of Usenet discussion groups by exposing spam and other off-topic postings.

Socially Embedded Processes

Having computers imitating socially embedded processes, such as communication, collaboration, cooperation, negotiation, or recommendation, always raises a couple of important issues that have to be dealt with. Benefit is gained through automation since protocols and procedures can be handled more efficiently by computers compared to their human counterparts. However, if a social process is reduced to the exchange of tokens according to a protocol, the remaining process does not capture the *social* nature of the process involving mutual commitments, being under obligation, being responsible, etc. (Lueg & Müller 1996). The conceptualization of social processes as basically "technical" processes is in the tradition of the "rationalistic" perspective (Winograd & Flores 1986).

Put in a nutshell, the rationalistic perspective assumes that the world can be described objectively and that optimal (rational) solutions to problems can be deduced from these objective descriptions. Implications of the rationalistic perspective in the information filtering context are manifold (Lueg & Pfeifer 1997). For example, it is assumed that the "content" of a document can be observer-independently estimated on the basis of its representation. Also, it is assumed that "interest" can be estimated independently from the actual situation the recipient of information is involved in. Accordingly, it is assumed that ratings given by a particular person in a specific context can be appropriately represented in numeric ratings and that it makes sense to de-contextualize these ratings.

Regarding recommendations, the social context of a recommendation is abstracted away from its social embedding; the recommendation is de-contextualized. Apparently, most current approaches to collaborative filtering are in the tradition of the rationalistic perspective.

Situated Cognition

From a cognitive science and situated cognition perspective, the so-called "rationalistic" perspective does not appropriately capture human cognitive phenomenons, such as cognition, knowledge, or behavior. Moreover, the rationalistic perspective does not provide an appropriate explanation for the notion of interest which is of outstanding relevance in the information filtering context. Contrary to the rationalistic perspective, which views human cognition as data-processing and behavior as being largely predetermined by plans, the situated cognition perspective suggests to view cognition, knowledge, and behavior as being fundamentally *situated*: cognition and knowledge are emergent properties of the interaction of an individual with its environment, i.e., its current situation (thus, the term "situatedness"). Cognition cannot be reduced to internal "data-processing", it cannot be "de-contextualized" into a set of abstract descriptions (Suchman 1987; Clancey 1997). One important implication of situatedness is that the way a human interacts with a situation continuously changes based on his or her experience. Accordingly, we propose to view interest as being dynamically generated: interest is an emergent property of the interaction of an individual with an "information situation".

Various approaches to find out about interest from different disciplines, such as psychology, information science, or computer science, can be found in the literature. Research on the notion of interest indicates that it is hard to determine why a specific document has actually been selected. Experiments (e.g., Lantz 1993; Mock 1996) have revealed that explanations of why a document was chosen for reading, or why it was found to be interesting varied and changed over time. The same result has been obtained when the subjects were asked about their initial information need. Situatedness explains why it is so hard to describe an information need. Information needs cannot be reduced to internal information processes alone, but require interaction with the current situation. Situational factors other than just the topical content of a selected document influence the relevance judgment. Factors influencing the judgment are any factors that the users bring into the situation, such as experience, background, knowledge level, beliefs, and personal preferences (Barry 1994). Also, the user's judgment is influenced by the user's purpose, the user's expectation, the relevance of references, and future time savings (Su 1994). Accordingly, divergences between professional research judgments of relevance and precision, and actual user judgments have been reported in the literature (Su 1994).

Situated Information Filtering

In general, the situated perspective applied to information filtering suggests that the goal is not to automate but rather to support information seeking processes in order to allow for situatedness and the peculiarities of human cognition. In an individual-based information filtering project, this perspective has lead to the development of `spynews` (Lueg 1997), a newsreader supporting users in acting situated while browsing Usenet newsgroups. Instead of trying to find out about user interests as in traditional approaches, the newsreader monitors the user's newsreading behavior and uses a discussion-oriented approach to find out in what he or she is *not* interested. This allows `spynews` to filter uninteresting discussions in order to help the user focus on potentially interesting discussions.

It's a peculiarity of `spynews` that no model of interests is being constructed to draw inferences about the user's interests. Also, no content analysis of selected documents (Usenet articles in this context) is performed to find out why particular documents have been selected. `spynews` only reflects the user's behavior by gradually fading out uninteresting discussions. Since no model of interests is constructed, the situated information filtering approach avoids the abstraction problem that occurs when documents or user interests are formally described and compiled to profiles (Lueg 1998).

`spynews` has been implemented as an augmentation to the state-of-the-art `Knews`[1] newsreader. Preliminary tests with experienced Usenet users are encouraging. Additional extended user tests are under preparation in order to evaluate the benefits of this particular approach. So far, the `spynews` newsreader only tries to find out about in which discussions the user is *not* interested in. We extend the newsreader to provide additional hints to *interesting* discussions. In order to account for situatedness, these hints will also be based on the user's browsing behavior only. Examples for user actions that can reasonably be interpreted as indicators of interest are reading a particular discussion -partly or completely- or posting a followup article. Also, external user actions, such as sending email to a participant of a discussion or saving an article might be interpreted as indicators for interest in a discussion. However, *all* user actions are only *weak* indicators, since there are many other explanations that are equally plausible: The participation might be nothing more than a final statement and an (interesting) discussion might be ignored due to too much time pressure, or the user might want to think more about a topic before entering into the discussion, etc. (Lueg & Pfeifer 1997).

So far, our research on situated information filtering has been focusing on *individuals.* Applied to the *community-based* collaborative filtering approach, the situated perspective suggests that the social embedding of recommendations should be considered more seriously. A personal recommendation does not only depend on the particular situation of the recommender but also on the relation between the recommender and the recipient of the recommendation. Of course, editors of recommendations in journals hardly know their customers personally but they always have at least an idea of the target audience. It is yet unclear how this social embedding might be utilized in a general recommendation context. A practical example from the Usenet domain might help illustrate the social embedding of recommendations.

Discussions within Usenet on detecting "interesting discussions" showed that it is typically not only the topic of a discussion that influences whether the discussion is interesting or not. In addition, it is of outstanding relevance which persons contribute to a discussion. Although most people participating in the global, distributed conferencing system Usenet news do not know each other personally, one can observe a kind of emergent regard among the participants of a discussion group concerning the opinion of particular persons and the way they articulate their opinions. Interest in the people's opinions might even outvote a less interesting topic. The situated perspective suggests that exploiting this particular social embedding for filtering purposes requires a careful investigation of the issue. In what follows, we discuss several related issues.

Future Research

Further research on information filtering and information overload situations is related to exploring the foundations for self-organizing "preference" networks, and investigating the usability of social filtering for spam-fighting and exposing clear off-topic postings. In the following, these issues are discussed in more detail.

Networks of Trusted Users Reports on experiences with Grouplens (Resnick *et al.* 1994), a collaborative filtering system for Usenet articles,

[1] `http://www.student.nada.kth.se/~su95-kjo/knews.html`

have shown that user acceptance is crucial especially at the beginning of a new collaborative filtering service since a critical mass of ratings is required for a working system (Miller, Riedl, & Konstan 1998). It has been argued that a kind of formal or implicit market system might be necessary to gain a sufficient number of ratings and to compensate those who consume ratings but do not provide ratings themselves (Konstan *et al.* 1997; Avery & Zeckhauser 1997). We investigate the development of tools supporting users in exchanging particular preferences with selected trusted people sharing interests.

Finding out about users sharing interests is a hot topic in collaborative filtering. The idea is that the preferences of one user with particular interests can be used as recommendations for other users with similar interests. However, if Usenet participants are viewed as situated agents that are embedded in a particular social environment (Usenet is best viewed as a virtual community), computing and comparing profiles in oder to find out about shared interests turns out to be obsolete since people automatically find out about other people sharing their interests by participating in Usenet discussions. Trust in the judgments of others and regard to the opinions of others emerges the same way.

Familiarity with other Usenet participants is an emergent property of participating in Usenet discussions. This familiarity might be used as a basis for self-organizing networks of people trusting each other and exchanging profiles of likes and dislikes among them. We suspect that such a distributed network might provide sufficient social embedding to avoid the above mentioned motivational problems. Since Usenet itself is a self-organizing network of servers, chances are not too bad that such a preference network might be accepted within the Usenet community. Moreover, this distributed approach would avoid some of the resource problems that centralized approaches, such as the Grouplens system (Konstan *et al.* 1997), exhibit. Also, security problems would be less serious since interest profiles are only exchanged among users knowing and trusting each other.

Exposing Spam and other off-topic postings Net abuse is a hot topic within the global Usenet community. A collaborative filtering approach might turn out to be a powerful tool to fight spam and to expose off-topic postings. Spam[2] denotes the flooding of Usenet newsgroups with commercial advertisements. Negative effects of flooding newsgroups with spam (certain newsgroups exhibit up to ninety percent spam) are manifold. Users already using spammed newsgroups are driven away since they increasingly have problems to detect new articles among uninteresting spam. While technically experienced participants may cope with spam by using sophisticated killfiles, new users not equipped with killfiles are kept away from spammed newsgroups.

Since hardly any participant in a newsgroup is interested in spam, keeping a newsgroup spam-free might provide enough motivation for the participants to provide ratings for a collaborative spam-filtering system. Besides having (seemingly) spam-free newsgroups, such a collaborative spam-fighting experiment would also provide valuable insights into the relation of varying interests among the participants of newsgroups and the motivational problems exhibited by traditional collaborative filtering systems. If such as collaborative spam-fighting system experiences a significantly better user acceptance than a traditional system, varying interests among the users of newsgroups could be identified as a reason for motivational problems keeping users from providing ratings in traditional collaborative filtering systems.

Summary

Social filtering experiments in the Usenet domain have turned out to be less successful than expected. Motivational problems seem to prevent people from providing a sufficient number of ratings in order to bootstrap a successful collaborative filtering process. In this paper, we have argued that this failure may be due to not sufficiently considering situatedness and a lack of social embedding. Based on our work on a situated filtering approach focusing on individuals, we have pointed out various issues that should be treated more carefully in order to reach a higher degree of user acceptance. In addition, experiences with Usenet suggest that a self-organizing network of people exchanging preferences might be an alternative to centralized collaborative filtering approaches. Also, a collaborative filtering approach might turn out to be a powerful tool to fight net abuse, such as commercial advertisements flooding newsgroups (also called spam). In addition, collaborative filtering can help to maintain the focus of newsgroups by exposing off-topic postings.

Acknowledgments

The authors would like to thank Michelle Hoyle and Ralf Salomon for their help, and Rolf Pfeifer for a stimulating research environment. We are grateful to Karl-Johan Johnsson for the permission to use and to modify his `Knews` newsreader for our Usenet experiments.

[2] `http://spam.ohww.norman.ok.us/default.htm`

References

Avery, C., and Zeckhauser, R. 1997. Recommender systems for evaluating computer messages. *Communications of the ACM* 40(3):88–89.

Barry, C. 1994. User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science (JASIS)* 45(3):149–159.

Brewer, R., and Johnson, P. 1994. Collaborative classification and evaluation of Usenet. Technical Report CSDL-TR-93-13, Department of Information and Computer Sciences, University of Hawaii.

Clancey, W. J. 1997. *Situated Cognition. On Human Knowledge and Computer Representations.* Cambridge University Press.

Denning, P. J. 1982. Electronic junk. *Communications of the ACM* 25(3):163–165.

Goldberg, D.; Nichols, D.; Oki, B. M.; and Terry, D. 1992. Using collaborative filtering to weave an information tapestry. *Communications of the ACM* 35(12):61–69.

Konstan, J. A.; Miller, B. N.; Maltz, D.; Herlocker, J. L.; Gordon, L. R.; and Riedl, J. 1997. Applying collaborative filtering to Usenet news. *Communications of the ACM* 40(3):77–87.

Lantz, A. 1993. How do experienced users of the system Usenet news select their information? Technical report, Department of Computer and Systems Science, University of Stockholm.

Lueg, C., and Müller, M. 1996. Cooperative systems: The right direction? In *Proceedings of the Second International Conference on the Design of Cooperative Systems*, 315–329.

Lueg, C., and Pfeifer, R. 1997. Cognition, situatedness, and situated design. In Marsh, J. P.; Nehaniv, C. L.; and Gorayska, B., eds., *Proceedings of the Second International Conference on Cognitive Technology*, 124–135. IEEE Computer Society.

Lueg, C. 1997. An adaptive Usenet interface supporting situated actions. In *Proceedings of the 3rd ERCIM Workshop on User Interfaces for All.*

Lueg, C. 1998. Supporting situated actions in high volume conversational data situations. In *Proceedings of the Annual ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'98).* ACM Press.

Miller, B. N.; Riedl, J. T.; and Konstan, J. A. 1998. Experiments with GroupLens: Making Usenet useful again. In *Proceedings of the 1997 USENIX Winter Technical Conference.*

Mock, K. J. 1996. Hybrid hill-climbing and knowledge-based techniques for intelligent news filtering. In *Proceedings of the National Conference on Artificial Intelligence (AAAI'96).* Menlo Park, California: AAAI Press.

Resnick, P., and Varian, H. R. 1997. Recommender systems. *Communications of the ACM* 40(3):56–58.

Resnick, P.; Iacovou, N.; Suchak, M.; Berstrom, P.; and Riedl, J. 1994. GroupLens: An open architecture for collaborative filtering of netnews. In Furuta, R., and Neuwirth, C., eds., *Proceedings of the International Conference on Computer Supported Cooperative Work (CSCW'94)*, 175–186. ACM Press.

Shardanand, U., and Maes, P. 1995. Social information filtering: Algorithms for automating "word of the mouth". In Irvin Katz et al., ed., *Proceedings of the Annual ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'95)*, 210–217. ACM Press.

Su, L. 1994. The relevance of recall and precision in user evaluation. *Journal of the American Society for Information Science (JASIS)* 45(3):207–217.

Suchman, L. 1987. *Plans and situated actions - The Problem of Human-Machine Communication.* Cambridge University Press.

Terveen, L.; Hill, W.; Amento, B.; McDonald, D.; and Creter, J. 1997a. PHOAKS: A system for sharing recommendations. *Communications of the ACM* 40(3):59–62.

Terveen, L.; Hill, W. C.; Amento, B.; McDonald, D.; and Creter, J. 1997b. Building task-specific interfaces to high volume conversational data. In *Proceedings of the Annual ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'97) Conference Proceedings*, 226–233. ACM Press.

Winograd, T., and Flores, F. 1986. *Understanding Computers and Cognition: A New Foundation for Design.* Ablex Publishing Corporation.

Knowledge Pump: Community-centered Collaborative Filtering

Natalie Glance, Damián Arregui and Manfred Dardenne

Xerox Research Centre Europe, Grenoble Laboratory

October 27, 1997

Abstract

This article proposes an information technology system we call the Knowledge Pump for connecting and supporting electronic repositories and networked communities. At the time of writing, we have a working prototype that we'll be ready to deploy soon within a first group of users. Our first goal is to achieve proof-of-principle: to show that community-centered collaborative recommendation can indeed support knowledge sharing and improve community awareness and development.

1 Introduction

This article proposes an information technology system we call the Knowledge Pump for connecting and supporting electronic repositories and networked communities. Our objectives are two-fold. The first is to facilitate getting the right information to the right people in a timely fashion. The second is to map community networks and repository content. These goals are complementary because the community and repository maps help channel the flow of information while the patterns inferred from information flow help refine the maps.

The aspect of Knowledge Pump on which we focus primarily here is its distribution capability. In particular, our first goal is to help communities, defined by their common interests and practices, more effectively and more efficiently share knowledge, be it in the form of must-read documents or new ways to get work done. We introduce a technique we call community-centered collaborative filtering. This technique combines statistical algorithms and heuristic rules with a community bias to guide the distribution of information based on explicit and implicit recommendations.

In the next section, we describe our first implementation of the Pump, and in Section 3 we conclude with a summary and outlook. A more complete elaboration of the system can be found in (1).

2 Implementation

In this section we present our first implementation of the Knowledge Pump. At the time of writing, we have a working prototype that we'll be ready to deploy soon within a first group of users.

2.1 Technologies and architecture

We had a few, basic initial design requirements: portability, ease of use and immediate value. Portability means one code set, all platforms, and suggested building something riding on top of the Web for a first implementation. Effectively, this pointed to Java, since HTML and scripting languages alone are too limiting. Portability also means not touching the browser: no plug-ins, for example, and no browser-specific capabilities, like cookies. In addition to that, we brought up an Apache HTTP server to access the system HTML pages and Java applets, and an mSQL shareware database engine to keep the whole system persistant data about users, documents, etc.

The Pump is implemented as a client-server system. The client is written in Java and runs off a Web browser. It talks to the Knowledge Pump server, also written in Java, which is responsible for a number of functions. The KP server provides an interface to system administration, periodically runs the community-centered collaborative filtering algorithm and builds the "What's recommended?" pages for each user. The pages are then saved and delivered to the user via the HTTP server. The database is accessed via yet another server.

In choosing to connect first to the Web, we've joined a well-populated playing field of WWW filtering efforts. However, we doubt that collaborative filtering over such a large domain can work well in an organizational setting: too many pages and too few reviewers. Our real goal is to connect to repositories which have Web interfaces. We're designing a plug-and-play interface to the Pump so that with minor modifications, any repository (Web front-end or no) can connect transparently with the Knowledge Pump. In the meantime, the Pump connects with any repository with a Web front-end in a less transparent, but still useful way. Currently, we are testing the plug-and-play principle by connecting the Pump with an electronic repository of scanned journal articles, (2).

2.2 Functionality: document management and recommendation

The user's interface to the Pump, shown in Figure 1 is through a small palette of functions. The *bookmark* and *search* functions are basic document management capabilities provided by the Pump. *Bookmark* allows the user to save a pointer to any Web page. (S)he rates the document on a five-point scale from "irrelevant" to "one of my favorites," and optionally types in some comments. The user also classifies the documents into any of a number of the listed communities. Finally, the user can save the pointer as "private" – for his/her eyes only – or as "public."

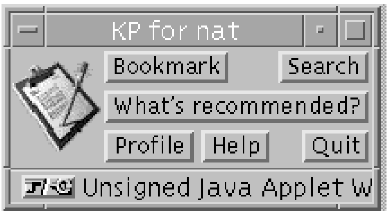


Figure 1: The user's palette of controls.

Complementary to the bookmark function is the *search* function, which allows the user to search for bookmarks classified into any number of domains according to date, title, author, rating, reviewer and private vs. public classification. The Pump delivers the search results as an HTML document to the Web browser. For each pointer satisfying the search criteria, the page includes the predicted or actual rating and provides a hyperlink to the comments associated with the pointer.

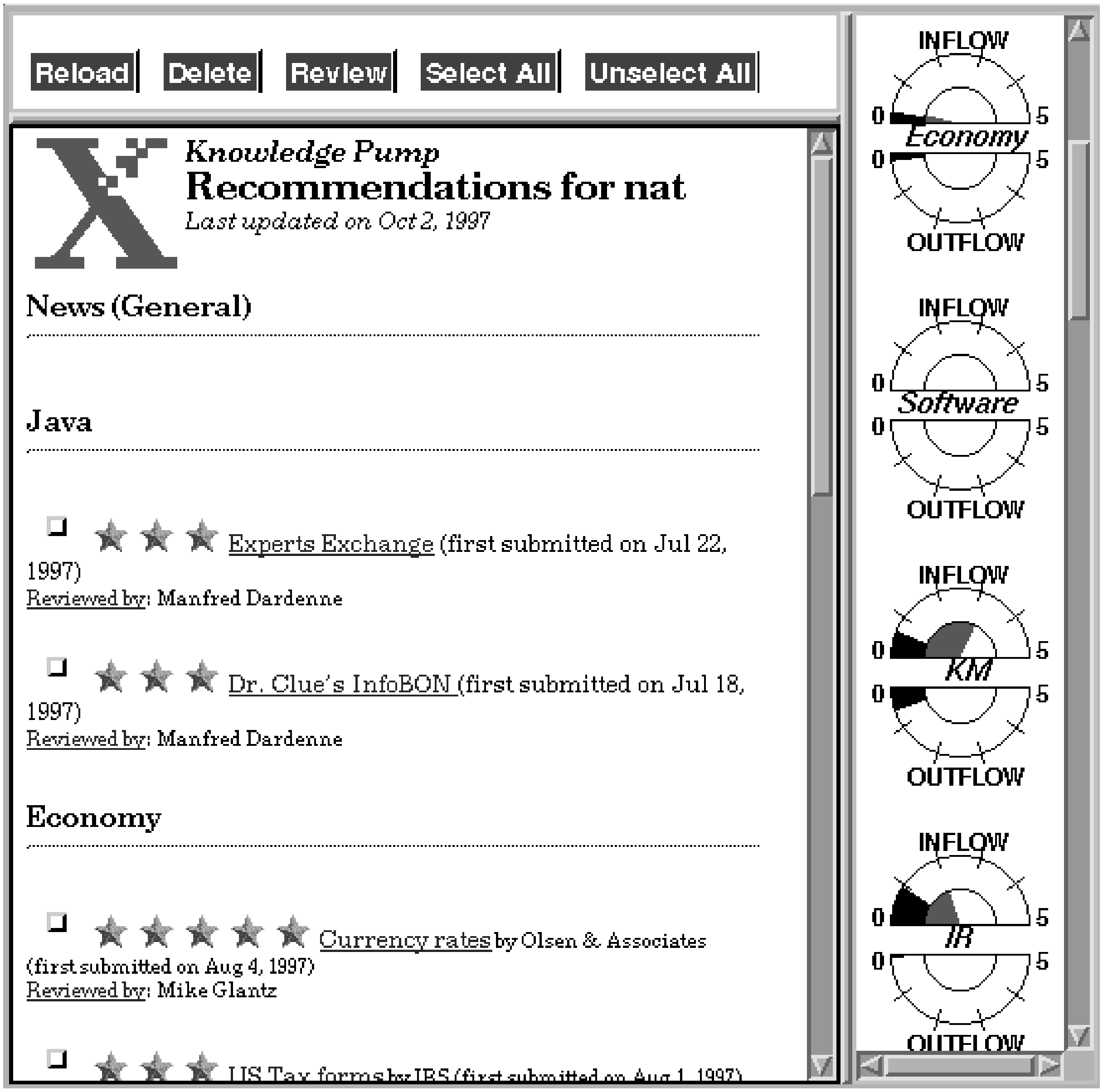
The *profile* function allows the user to enter or modify his/her personal profile. Here the user selects his/her "advisors" from the list of Knowledge Pumpers. Advisors refer to people whose judgement the user particularly respects – this list is used by the community-centered collaborative filtering mechanism described further in Section 2.3. The user also selects any number of domains of interests from the hierarchy of "communities." Recommendations by the Pump to the user will be sorted according to this identified set of domains.

The *What's recommended?* function brings up the Pump's most recent personalized list of recommendations sorted by category as an HTML document in the Web browser. An example recommendations page is shown in Figure 2. If the user keeps the recommendation page open in the Web browser, the Pump periodically and automatically updates it. In the current implementation, each recommendation takes the form of a pointer to a URL. The Pump only recommends items which the user has not seen before (not to the Pump's knowledge, at least) and for each item, displays the Pump's prediction of the user's interest as a number of stars, lists the names of all reviewers, and provides a link to their comments. The user can prune the recommendations page by deleting entries and can also review items directly from the page.

The panel to the right of the recommendations page contains a set of *gauges*, as we call them. These gauges reflect the activity level of the Pump. There is a gauge displayed for each community to which the user belongs. The INFLOW half of the dial indicates how many recommendations are flowing in per person per week for the community. In black is the community average; in red is the individual inflow. The OUTFLOW part of the dial indicates how many recommended links are being followed per person per week. Once again, black represents the community average; red, the individual outflow. The gauges give feedback on how the average level of activity in a community fluctuates over time and give users a feel for how their level of participation compares with community averages.

2.3 Community-centered collaborative filtering

The Knowledge Pump uses what we call community-centered collaborative filtering to predict a user's level of interest for unread items in each of the users' domains of interest. This mechanism combines elements of social and content-based filtering.



Knowledge Pump

Recommendations for nat

Last updated on Oct 2, 1997

News (General)

Java

□ ★★★ Experts Exchange (first submitted on Jul 22, 1997)
Reviewed by: Manfred Dardenne

□ ★★★ Dr. Clue's InfoBON (first submitted on Jul 18, 1997)
Reviewed by: Manfred Dardenne

Economy

□ ★★★★★ Currency rates by Olsen & Associates (first submitted on Aug 4, 1997)
Reviewed by: Mike Glantz

□ ★★★ US Tax forms by IRS (first submitted on Aug 1, 1997)

Figure 2: An example of "What's recommended?" by the Pump.

For the moment we rely on recommenders to classify items into a commonly agreed upon classification scheme. This could be complemented down-the-line by automatic categorization via statistical classification algorithms.

The second layer of social filtering – matching items to people by first matching people to each other – lies on top of the initial classification by domain. It's important to filter by content first and by social relationships second because similarities among people tend to vary greatly across different domains. For example, the authors of this article have similar rankings of the most influential knowledge management gurus, but wildly different opinions concerning the best guitar players alive. Social filtering over all domains at once tends to wash out the differences in people's similarities to each other.

Social filtering via automated collaborative filtering is based on the premise that information concerning personal relationships are not necessary. In principle, we agree, because once an automated collaborative filter has collected enough information about its users, it can work very well. In practice, however, automated collaborative filters suffer from the cold-start problem: without large amounts of usage data, they work very poorly, which discourages the usage that would overcome this lack.

In contrast, in community-centered collaborative filtering, the collaborative filter is bootstrapped by the partial view of the social network constructed from user-input lists of "advisors" – people whose opinion users particularly trust. Bootstrapping the system in this way allows the collaborative filter to perform well from the start, weighting higher the opinions of his/her most trusted contacts when predicting the user's opinion on items. Over time, as more usage data is collected, the weight given to automated (statistical) portion of the collaborative filter can be increased relative to the weight given to advisors' ratings.

Statistical algorithms can then mine the usage data to automatically refine the Pump's view of the social network and visualize it for the users. This sets up a feedback loop between users and the collaborative filter: on their end, users collectively (re-)describe the social network; on its end, the Pump automatically refines and visualizes the social and community maps from usage data.

Item #	Alice	Bob	Chris	Dave
1	5	?	3	4
2	?	1	5	2
3	?	4	2	?
4	0	?	1	?
5	?	?	3	3

Figure 3: A sample user-item matrix of ratings.

From a mathematical standpoint, collaborative filtering within a given domain can be viewed as matrix filling, where the rows of the matrix are items recommended into the domain, the columns are the people who have reviewed an item in the domain, and the cells contain the ratings submitted. An example is shown in Figure 3.

The prediction algorithm used by the Pump is a weighted sum of three components:

- the average population-wide rating;
- the average over advisors' ratings;
- the correlation-weighted sum of all ratings.

The first two components are straight-forward and are very important when ratings are very sparse, for example, when the system is first deployed. The second component uses the elements of the social network revealed from user-input lists of advisors.

The third component is a standard automated collaborative filter (see (3), for example), which can be implemented in any of a number of ways. Our implementation first calculates person-person correlations from previous recommendations. These correlations indicate how much two reviewers tend to agree with each other on the items they both rated.

The collaborative filter automatically weighs more heavily the ratings of users that historically tend to agree with the user in question and discounts the ratings of those that tend to disagree. It is most effective when the user-item matrix is densely filled.

Currently, we use heuristics to combine the three components into one prediction. The heuristics take into account how long the system has been in place and the confidence level of each of the three elements. The confidence level is simply an ad-hoc estimate based on the density of ratings in the respective three populations. Once we have a user base established, we'll be able to test the effectiveness of our approach and of the confidence level estimates and refine them for future use.

One last element of community-centered collaborative filtering is related to the user interface: users see the names of the people who have recommended an item and can read the publicly-available comments. This makes the boundaries between communities more permeable. A user can classify an item into any domain, regardless of whether (s)he considers him/herself as a member of that community. Thus, members of a community can receive recommendations from people outside the community. Over time, the person can explicitly join the community by changing his/her profile or may become a de facto participant in the minds of its members.

3 Summary and outlook

As we discussed earlier in this article, we believe that the key to sucessful knowledge sharing is focusing on the community. We've implemented a first version of the Knowledge Pump that attempts to leverage community currency in the form of reputation, trust and reciprocity to create incentives for sharing recommendations. At the heart of the Pump is a recommendation distribution mechanism we call community-centered collaborative filtering. This mechanism matches items to people by first matching people to each other, giving extra weight to trusted advisors.

The implementation of the Knowledge Pump as described in the previous section is a work in progress. Our first goal is to achieve proof-of-principle: to show that community-centered collaborative recommendation can indeed support knowledge sharing and improve community awareness and development. This first prototype is intended to provide the minimal set of functionality sufficient to make it acceptable for use within an environment of early adopters.

However, understanding the environments in which the Pump could be used will be vital in order to tailor its functionalities and create incentives for use. For something like the Knowledge Pump to successfully support the flow and use of knowledge in organizations, it will have to become a seamless part of the way people do their work. The social aspects of use are perhaps the most fascinating and the most challenging.

4 Acknowledgments

The authors thank Daniele Pagani and Dan Holtshouse for inspiring this work and for their continued support. We are also grateful to Stefania Castellani, Antonietta Grasso, David Hull, Francois Pacull, Remo Pareschi and our other colleagues at XRCE for their suggestions and feedback.

References

[1] N. Glance, D. Arregui, and M. Dardenne, “Knowledge pump: Supporting the flow and use of knowledge,” in *Information Technology for Knowledge Management* (U. Borghoff and R. Pareschi, eds.), ch. 3, Springer-Verlag, 1998.

[2] URL. Calliope: http://www.xrce.xerox.com/ats/digilib/calliope/.

[3] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, “Grouplens: An open architecture for collaborative filtering of netnews,” in *Proceedings of the Conference on Computer-Supported Cooperative Work*, (Chapel Hill, NC), pp. 175–186, ACM, 1994.

Social Affordances and Implicit Ratings for Social Filtering on the Web

Rob Procter[1] and Andy McKinlay[2]

[1] Department of Computer Science, Edinburgh University
Edinburgh EH9 3JZ, Scotland
[2] Department of Psychology, Edinburgh University
Edinburgh EH8 9JZ, Scotland

Introduction

The theme of this paper is exploring ways of extending web browsing environments to facilitate the sharing of information pertaining to document quality amongst communities of users on the Web. Amongst the issues it raises for discussion are:

- sources of rating and recommendation data,
- the context of ratings and recommendations,
- real and virtual groups, and
- privacy and accessibility.

Much of the current work on social filtering tools for the Web has focused on so-called explicit methods, i.e., where the rater annotates a document or (more simply) inputs a rating value (e.g., [12]). One drawback of this approach is that it calls for extra effort on the part of the rater, whilst failing to provide an equally immediate benefit [5]. In contrast, implicit methods require no extra effort on the part of the rater, but have the disadvantage that the rating information provided has lower value. Some tools have attempted to find some middle ground between explicit and implicit approaches [7]. Our interest here lies in exploring how implicit approaches might be improved to provide rating information and higher value and relevance.

Social Filtering

Social, or collaborative filtering is an emerging technique for dealing with overload in information environments (i.e., systems for the production, dissemination and consumption of information). One widely explored technique for social filtering is based upon ratings and recommendations which are supplied by readers and disseminated for the guidance of new readers. One well-known example of this technique is the GroupLens system, which has been implemented for filtering Usenet news postings [10].

There are several major difficulties with any kind of reader ratings-based approach. These include:

- the cost to readers of generating ratings [7], and
- how readers become trusted (as raters) and learn to value and trust the ratings of others.

In the GroupLens project, this first issue is addressed by the empirically verified observation that the time spent reading a Usenet posting is itself an implicit measure of the reader's valuation of it [10]. In principle, therefore, Usenet ratings can be gathered cost-free.

In conventional communities, the issue of trust is resolved through community and social interaction: people learn of one another's interests and experiences and reputations develop which serve to establish the value of the ratings currency. Though Usenet's communities are virtual, it nevertheless has a strong strong community orientation [23]. However, in seeking to determine the value of ratings, GroupLens side-steps the problem of trust as a social fact altogether. Instead, it uses a statistically-based predictive algorithm which establishes an historical match between news group readers' ratings and then uses this to determine the significance of ratings for new postings. In GroupLens, therefore, trust is merely a computed relationship between readers rather than a consciously evolved and acknowledged social fact.

Nowhere are the problems of information overload more evident than on the Web. People often have great difficulty in finding information of value. Already, commercial recommendation services have become widely available (e.g. Yahoo). Some services address broader measures of quality of service at the site level by collecting and publishing statistics of server response times, site maintenance standards (e.g., proportion of broken links), and also more subjective measures such as aesthetic design. These are all valuable resources for the information seeker, but they do not address all their needs. Probably one of the most common ways still of obtaining recommendations for Web pages is a URL in an email message from friend or colleague, or in a UseNet posting. Not surprisingly, therefore, tools of integrating email and UseNet with the Web have attracted some interest [2, 12].

In the following sections, we discuss ways in which Web users' behaviour may provide some of what is presently missing from Web rating and recommendations services.

Social Filtering and the Web

Whilst GroupLens successfully achieves the goal of relevant implicit ratings, are numerous reasons why its approach cannot be simply transplanted to other environments such as the Web. As a genre, Usenet has a number of attributes which are essential for social filtering. First, Usenet is founded upon the concept of community: news groups are not just thematic devices for identifying content, they also provide users with the experience of group membership which is crucial to collaborative activity [3, 23]. Second, news groups are an interactive genre where information is both produced and consumed within the news group community. The news group provides both the context for matching ratings information and the experience of community which makes this information relevant and meaningful to recipients.

In contrast, the Web is founded upon an abstract information model, rather than upon community and collaboration. Though there is a sense of place in the Web, it is place as in Web site, rather than place as in community. Furthermore, the Web is inherently less interactive: processes of information production and consumption are more clearly separated. Unlike Usenet, the Web does not incorporate an explicit model of community and interaction. It is primarily intended as a vehicle for information distribution and foraging.

To summarise: Usenet's metaphor for information environments is the group discussion — i.e., it offers only minimal simple structuring devices (the newsgroup and thread), but compensates for this with with richer interactivity. In contrast, the Web's metaphor is the library — i.e., it offers a relatively sophisticated structuring device for information, but only limited interactivity.

We now consider ways in which community and interactivity on the Web can be enhanced and how this may contribute to effective social filtering.

Social Affordances and the Web

In user interface design, affordance is defined as making the potential for action visible"... a technical term that refers to the properties of objects — what sorts of manipulations and operations can be done to a particular object" [13]. Its application as a design principle is ubiquitous in the graphical user interface; for example, the rendering of screen button images as objects with depth, affords the action of pushing. By analogy, social affordance can be defined as the "... making the potential for social (inter)action visible".

Physical environments are rich in social afforances. Shared spaces afford knowledge about what activities are being performed and who are performing them. They also afford knowledge about how the activities are being performed and the artefacts employed. Physical environments afford social learning — i.e., the use of others as social tools.

Physical workspaces are rich in social affordances which help their occupants remain aware of others are doing. In turn, this awareness facilitates collaboration of both a formal and informal nature. For example, in the conventional library space library users may gain helpful clues about where to search for a particular item, or they may see a colleague who may be able to give assistance. Similarly, conventional information artifacts such as library books may be sources of useful rating and recommendation information: Twidale and Nichols, for example, cite the instance of the frequently borrowed, well-thumbed book [22]. The Web, like many other forms of digital information resources, lacks these social affordances. People's activities become less publically available through being screen-based, and network accessibility reduces the need for performing these activities in public places.

The different character and properties of digital information artifacts also has important implications for social affordances. Indeed, within in modern media as a whole, processes of social demassification — i.e., the disaggregation of large social units into smaller groups — are very much in evidence [1]. Traditional information artifacts like newspapers are useful not simply because they provide information to the individual reader, but also because this is 'social information' — i.e., common to a broad readership. Part of its value is that everyone is reading it [1]. In contrast, on the Web, no one knows just who is reading what.

The general question then is how social affordances can be incorporated into Web environments [16]. There are two specific questions:

1. how can readers' actvities be made available to one another?
2. how can use of Web artefacts be made (more) public?

The use of readers' Web bookmarks is one way of exploiting information artefacts in social Web filtering [17]. Shareable page annotations are another [4, 18]. Both of these approaches have the disadvantage that they require readers to make a specific effort to record their preferences (though this effort is mitigated by the fact that the bookmarker or annotator is acting for their own benefit). However, surveys of Web users provides evidence that they typically bookmark fewer than 50% of the pages they find interesting; bookmarks tend to be evidence of strong, rather than marginal interest, so they set a high threshold for recommendations [17].

In contrast to these explicit approaches, we propose examining what kinds information about Web page ratings may be inferred of more informal or implicit evidence of users' browsing behaviour and how this might be shared within groups of Web users. The value of past browsing patterns as a predictor for a user's current and future information needs has already been demonstrated [15]. We are interested in the value that an individual's or a group's browsing behaviour may have for other web users.

Social affordances of the Web

The Web provides its users with a shared information space. Typically, however, the sharing of the Web is experienced by its users as a problem rather than as an opportunity. Popular Web sites cause frustration when overloading creates excessive delays in page downloading. Also, compared with spaces such as Usenet, the Web is poorly structured: its boundaries and borders are not clearly defined.

As shareable artefacts, the anonymity of the process (both synchronously and asynchronously) of Web page sharing makes this unusable. The paradigm shift from one-to-many, broadcast information dissemination (e.g., Usenet) to one-to-one narrowcast dissemination (e.g., Web) is not only an inefficient use of network bandwidth, but undermines the sense of community that broadcast methods engender [1]. As a medium of information dissemination and exchange, the Web lacks the features that are characteristic of community, including [14]:

1. social interaction,
2. clearly defined group boundaries; and
3. a capacity for members to monitor each others' behaviour.

Explicit and Implicit Ratings

A problem with both explicit and implicit rating approaches is that of poverty of context: ratings should make their origins apparent [7]. Explicit approaches de-contextualise rating information — i.e., they assume that ratings have a value which is independent of the circumstances in which they were generated [11]. Implicit approaches, in contrast, might claim to be naturally contextual, but face significant problems in utilising this context in a way which is both informative and shareable.

Contexts for implicit ratings

There are innumerable contexts of use, but the problem of contextualising Web users' behaviour can be simplified by considering ways in which the space of contexts may be partitioned. We suggest that the following three dimensions are of particular relevance for ratings systems:

1. people — who is doing the rating,
2. documents — the patterns of access, and
3. time — when the documents were read.

People

Who is doing the rating is obviously important for assessing its value and relevance. Scientific journals take care to assemble editorial boards from recognised experts in the field and prominently display their names. Electronic journals have, in the past, suffered from credibility problems because they failed to convince their audience that their quality controls matched those of more conventional journals [6]. When searching for quality documents on the Web, a good strategy might be to copy the browsing patterns of an acknowledged expert in one's field:

> "If I have identified a person or an institution who does excellent work in my area, it is advisable to follow the activities of the person or institution to the extent which is manageable to me." [20]

To do thus, however, might be unacceptable to the expert. Indeed, there are a number of reasons why it would be preferable to tracking browsing behaviour at the group level, rather than at the level of the individual. Preferably, this group should consist of similar people [7]. The question for the Web is how might such groups identify themselves? In Usenet, group members define themselves through their domain of interest. This is often thematically-based, and such groups are virtual rather than real. However, there are newsgroups whose membership is defined through looser affiliations such as organisation or location.[3] One reason that groups defined by organisational affiliation have a locus of interest, is that it pays for members of an organisation to be informed about one anothers' activities.

[3] This locality is often reinforced by restricted access.

Documents

In most existing ratings systems, the unit of analysis is the single document. In GroupLens, for example, it is the individual posting which is rated. Yet the individual posting may appear in the context of a group of postings, i.e., a thread. In such cases, it is legimate to ask whether the rating of an individual document is meaningful. If not, then the question is how might the document's context as one of a group of documents consulted in a sequence be taken into account.

Discourse analysis is an established technique for studying the conversational relationships between speakers' utterances [19]. Self-evidently, the significance of an individual utterance rests upon the context in which it occurs, rather than on its own content alone. A number of particular conversational relationships are cited by discourse analysis: coherent pairs, e.g., a question and answer pair, are said to be *sequentially accountable*; co-occurent (but not adjacent) utterances, are said to be *distributionally accountable*; a conversation is said to have *topical coherence* if the sequence of utterances are consistent.

The use of discourse analysis as a tool for analysing Web users' behaviour has been proposed by Jasper et al. [8]. They argue that different objects and links within a Web page may be said to be sequentially accountable to that page. Likewise, a set of pages which are reachable from a given page may be said to be distributionally accountable. Finally, topical coherence may be related with the content match across sets of Web pages within a specific time frame. We might explore also the notion of site coherence — i.e., the relationship between in-site and out-site page accesses.

Time

As a corollary of topical coherence, temporal coherence in Web access patterns may be defined as the degree of overlap between the page accesses amongst a group of Web users within a particular time frame — i.e., the synchronicity of page access patterns within the group. Time may be an important factor in determining the relevance of information about another group member's browsing behaviour. For example, it may serve as a cueing device for an event of collective interest. For example, the content of newspapers is determined largely by the criteria of timeliness.

Extracting Implicit Ratings From Web User Behaviour

The most basic of implict evidence of a Web page's value is simply the fact that it has been accessed. Of course, by itself, this may be unreliable, just as is the act of reading a Usenet posting. In the latter case, extra value can be extracted from the time that the reader spends reading the posting [10]. A similar approach to Web pages may also yield useful rating information.

Browsers can be configured to cache pages, so analysis of cache contents may provide not only information about the time spent reading a Web page, but also about page access patterns, which may be used to further enrich the raw page reading time data by adding document context. As an example, unlike Usenet postings, a particular Web page may become important as an anchor point in a sequence of linked page accesses. In this case, the important metric is not the time spent reading the page on a per visit basis, but the frequency of accesses [21]. This points to another important distinction between rating Usenet postings and Web pages: the latter may be rated not only for their nominal content, but also because of their navigational value — i.e., they serve as route markers for accessing other pages.

For reasons of page traffic reduction (amongst others), organisations which host multiple Web users often interpose a proxy machine between the user and external web sites. The proxy serves as a local cache of pages; new page requests are compared against the proxy contents and if a match is found, the page is retrieved from the proxy cache. The proxy cache is therefore one potential source of data for tracking Web access patterns and generating recommendations within a particular community of users — i.e., within a group context. The typical proxy cache replacement policy is based upon frequency of access: the time a page spends in the cache is therefore a measure of its collective recommendation rating.

Implicit in systems like Siteseer and proxy-based caches is the assumption that it is users' physical locality which establishes the natural boundaries of the recommending and filtering group. Providing Web users with virtual group proxies could provide an alternative approach to group membership definition which is more similar in concept to the Usenet newsgroup (WebCard adopts a similar approach (see [2]). Users could register with group proxies of their choice and so become members of virtual recommendation communities.

More sophisticated analysis of proxy page trafffic could be used to establish document context information. Following the principles of discourse analysis, Web document ratings could be weighted according to:

- nominal rating — aggregate page viewing time;
- frequency — the number of times a page is requested;
- sequential accountability — the number of objects and links within a page;
- distributional accountability — common sequences of page accesses within the community;
- sources — the number of times the page is identified as the beginning of a distributionally accountable sequence;
- topical coherence as measured by inter-document text similarity [15]; and
- temporal coherence — the temporal distribution of page accesses.

Navigation is an extremely important issue for Web use. It follows, that not only may it be valuable for Web users to have access to ratings of *collections* of documents (and Web sites), but also to 'good' route maps for navigation within document collections, and within Web sites. Such maps can be determined from aggregated distributional accountability analysis.

Issues in Making Web Browsing Public

The use of proxy caches as sources of recommendations raises several issues. One is how this information might be made available to group members. As the analogy here is the affordance of the shared workspace for observing what people in the group are doing; one strategy would be to make the proxy cache browsable. This in turn raises issues of privacy — people can take steps to limit the accessibility of their activities; and reciprocity — in the physical workplace, observers may themselves be observed.

By choosing the group as the unit of observability, rather than the individual, people's sensitivity over privacy may be reduced. People may nevertheless wish to retain some degree of control over what they make public and what they choose to keep private. The principle of the group provides the basis for applying access controls. At one extreme, people may choose not to be a member of such a group. Those that do join may be given a variety of control over privacy, which through the principle of reciprocity, establish not only what is made public to others, but also what they themselves are able to know about others.

Conclusions and Further Work

In this way, it may be possible to incorporate richer and more relevant notions of group into the Web environment which may, in turn, make implicit approaches more effective as resources for ratings and recommendations. Much work needs to be done on techniques for extracting context, particularly discourse analysis, to develop them and empirically verify their value.

There are many issues that we have left explored. A major one is to how the rating and recommendation information is made available to the user. Some possibilities include incorporating it within browsing as "recommendation enhanced" links and menus [7]. Another would be to circulate regular (e.g., daily) summaries and digests of the community's browsing. These, and many other issues, require further investigation and evaluation.

References

1. Brown, J. S. and Duguid, P. Borderline Issues: Social and Material Aspects of Design. Human-Computer Interaction, Vol. 9, No. 1, 1994.
2. Brown, M. Webcard = Mail + News + WWW. In Proceedings of CHI'97, (Extended Abstracts), ACM Press, Atlanta, March, 1997.
3. Donath, J. Identity and Deception in the Virtual Community. In Kollock, P. and Smith, M. (Eds.) Communities in Cyberspace: Perspectives on New Forms of Social Organization. Berkeley: University of California Press, 1997.
4. Greenberg, S. and Roseman, M. A Groupware Web Browser. In Proceedings of CSCW'96, ACM Press, Boston, November, 1996.
5. Grudin, J. Why CSCW applications fail: problems in the design and evaluation of organisational interfaces. In Proceedings of CSCW'88, ACM Press, New York, 1988, p. 85–93.
6. Harnad, S. Implementing Peer Review on the Net: Scientific Quality Control in Scholarly Electronic Journals. In Proceedings of the International Conference on Refereed Electronic Journals, University of Manitoba, Winnipeg, October, 1993.
7. Hill, W., Stead, L. Rosensteian, M. and Furnas, G. Recommending and Evaluating Choices in a Virual Community of Use. In Proceedings of CHI'95, ACM Press, Denver, May, 1995.
8. Jasper, J., Darin Ellis, R. and Wajahath, S. A Discourse Analysis of User Clickstreams on the Web. Submitted to Interacting with Computers.
9. Kamiyana, Roscheisen and Winograd. Grassroots: A system providing a uniform framework for communicating, structuring, sharing information, and organising people. In Proceedings of the Fifth International WWW Conference, Paris, May, 1996, p. 1157–74.
10. Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L. and Riedl, J. GroupLens: Collaborative Filtering for Usenet News. Communications of the ACM, March, 1997, p. 77–87.
11. Lueg, C. Social Filtering and Social Reality. In Proceedings of the Delos Workshop on Collaborative Filtering, Budapest, November, 1997 (this volume).
12. Maltz, D. and Ehrlich, K. Pointing the way: actice collaborative filtering. In Proceedings of CHI'95, ACM Press, Denver, May, 1995.
13. Norman, D. Turn Signals are the Facial Expressions of Automobiles. Reading, MA: Addison Wesley, 1992.
14. Ostrom, E. Governing the Commons: The Evolution of Institutions for Collective Action. New York: Cambridge University Press, 1990.
15. Pirolli, p., Pitkow, J. and Roa, R. Silk from a Sow's Ear: Extracting Usable Structures from the Web. In Proceedings of CHI'96, ACM Press, Vancouver, May, 1996.
16. Procter, Goldenberg, McKinlay and Davenport. Enhancing Community and Collaboration in the Virtual Library. In Peters, C. and Thanos, C. (Eds.) Proceedings of the First European Conference on Research and Advanced Technology for Digital Libraries, Pisa, September, 1997. Lecture Notes in Computer Science v. 1324. Springer-Verlag, p. 25–40.
17. Rucker, J. and Polanco, M. Personalized Navigation for the Web. Communications of the ACM, March, 1997, p. 73–75.
18. Schickler, Mazer and Brooks. Pan-Browser support for annotations and other meta-information on the World Wide Web. In proceedings of the Fifth International WWW Conference, Paris, May, 1996, p. 1063–74.
19. Schiffrin, D. Conversational Analysis. In Linguistics: The Cambridge Survey, Vol IV, Language: The Socio-cultural Context. Newmeyer, F. (Ed.), Cambridge: Cambridge University Press, 1989.
20. Sint, P. P. Institutional Rating in Everyday Life. In Proceedings of the Delos Workshop on Collaborative Filtering, Budapest, November, 1997 (this volume).
21. Tauscher, L. and Greenberg, S. Revisitation Patterns in World Wide Web Navigation. In Proceedings of CHI'97, ACM Press, Atlanta, March, 1997.

22. Twidale, T. and Nichols, M. In Proceedings of the 4th Uk Digital Libraries Conference, Milton Keynes, May, 1997.
23. Wellman, B. and Gulia, M. Net Surfers Don't Ride Alone: Virtual Communities as Communities. In Kollock, P. and Smith, M. (Eds.) Communities in Cyberspace: Perspectives on New Forms of Social Organization. Berkeley: University of California Press, 1997.

Networks of Language Processors: a language theoretic approach to filtering and cooperation

Erzsébet Csuhaj-Varjú *
Computer and Automation Research Institute
Hungarian Academy of Sciences
Kende u. 13-17
H-1111 Budapest
Hungary
E-mail: csuhaj@sztaki.hu

Abstract

Networks of language processors (NLP systems) is a collective term which has been introduced as a formal language theoretic framework for describing symbolic processing in highly (massively) parallel and distributed architectures. Roughly speaking, an NLP system consists of several language determining devices (language processors) which are located at nodes of a virtual graph (a network) and which rewrite strings and communicate them through the network. In this paper we briefly discuss the model and introduce a particular variant which can be considered as a formal model for collaborating agents which communicate with each other through a network and use recommendations for filtering information.

1 Introduction

One of the most challenging problems of current computer science is to develop sophisticated, highly reliable tools for supporting effective information dissemination and information search performed by users of computer networks. All who use Internet face similar questions every day: how to choose from the lot of information received and to be communicated, how to select the useful or the important ones from the multitude of the arriving messages. These and similar problems are frequently discussed in the case of groups of agents collaborating through networks and formulating and using recommendations for filtering information ([1]).

The solutions of these problems and the answers to these questions suppose having an elaborated semantic background, but to develop suitable and convenient software tools for supporting effective information filtering, also syntactic aspects have to be carefully studied.

A project, titled " Networks of Language Processors" started last year at the Research Group on Modelling Multi-Agent Systems, at the Computer and Automation Research Institute of the Hungarian Academy of Sciences, with the aim of describing at the pure syntactic level characteristics of the behaviour of agents and agent communities using a network for cooperation and communication and to offer tools for designing languages supporting collaborative text processing via networks.

The research is mainly based on tools of formal languages, a traditional area of theoretical computer science, and it is a continuation of investigations that have been done for years by an international team in a recent field of formal language theory called (parallel communicating) grammar systems ([9], [3]).

The developed framework is called networks of language processors (NLP systems). This collective term originally has been introduced as a formal language theoretic framework for describing

*Research supported by Hungarian Scientific Research Fund "OTKA" No. T 017 105

symbolic processing in highly (massively) parallel and distributed architectures ([6]). The model was strongly motivated by some known models and paradigms ([7]),[8],[11],[10]). Arguments for formulating such a concept were, among other things, the claim to provide reliable language theoretic support for networked computing, for social networks, for describing the behaviour of mainly locally connected processor arrays, and understanding the nature of massively parallel and distributed architectures, including ones with biological or other nature-motivated background.

NLP systems capture properties of some related notions from formal language theory: the test tube systems ([4]), language theoretic constructs for distributed architectures from DNA computing, parallel communicating grammar systems ([3]), models motivated by distributed and decentralized problem solving systems, grammar systems with WAVE-like communication, providing grammatical models of the so-called Logic Flow paradigm ([5]).

2 Networks of language processors

In the following we briefly describe the main characteristics of the framework. For further details and information the reader is referred to [6] and [2].

A network of language processors (an NLP system, for short) consists of several language determining devices (or mechanisms computing multisets of strings), called language processors. These form the components of the system.

Each language processor represents an agent which processes textual information and cooperates with the other ones by communicating information pieces. Every language processor is located at some node of a virtual graph (a network), moreover, there is no more than one language processor at each node.

The language processors of the NLP system operate on strings (on sets of strings or multisets of strings) by performing rewriting steps and communication steps, usually alternately. The strings can represent data and/or programs (the latter correspond to language theoretic operations in coded form, or to sets of rewriting rules); both kinds of them can be rewritten and communicated. The same string can be interpreted at different components in different manners: it can play the role of a piece of data at some component and that of a rewriting rule at some another one. Thus, the agents can modify the information they have available and they can communicate it to each other. This information can be textual data (strings representing data) or it can be some rule of information handling (program, operation code).

During the functioning of the system new agents can join the network and agents are allowed to leave the agents' community. This leads to a flexible, self-organizing topology of the network. Thus, creation of new components and deletion of some existing ones are allowed, which can be done both as a result of a rewriting step and/or a communication step.

The NLP system is functioning by changing its states (if the rewriting rule sets can be modified, then we use the term "configuration" instead of the term "state"). At any moment of time, the state of the network is described by the sets of string (multisets of strings) present at that moment at the components. Thus, at any moment of time the agents' community is represented by the collection of strings the agents have available at that moment.

At the beginning of functioning, each component of the NLP system is initialized by a language processor (a (finite) set of rewriting rules and the way of its application: for example, a production set of a grammar) and a finite set of initial strings, the axioms. These, together, form the initial state (or the initial configuration) of the system.

The change of the state of the NLP system can take place either by a rewriting step or by a communication step. By a rewriting step, some strings present at some component are rewritten according to the rewriting rule set and rewriting mode of the component (by the metarules in the case of changing the rewriting rules).

By a communication step, some strings (or copies of some strings) which are present at some component and satisfy some condition are communicated to one or more components. The target components are determined by a neighbourhood relation: each language processor is allowed to try to transmit strings only to its neighbours.

Thus, communication is realized through (mainly) local interactions among the components. (Clearly, in some very special cases, the neighbourhood relation makes a broadcast possible.)

The language processors in the network can work either in synchronized or in asynchronous manner.

During the functioning of the NLP system, the communication structure can dynamically vary or it can remain unchanged. The communicated strings are processed by the components which receive them. This can take place in various manners: for example, the arriving strings join the available string community of the component or they are concatenated to some of the strings present at the component.

The conditions for communication can be defined in several ways. One of the most important variants is that one where context conditions are imposed on the strings to check whether the current string can be communicated or not (for example, the existence of some kind of substrings of the string is tested). This, often, is given in the form of *filter* or *selector languages*, that should contain as an element the string to be sent or received. The components can have both an input and an output (entrance and exit) filter. Thus, each agent controls the information flow by using some selector mechanism in order to distinguish useful or important information from the arriving or sent messages. This is similar to what takes place in email systems: some messages have priorities to the other ones both in sending and receiving/reading.

Some components (agents) can have the same input and/or output filter: they form a *team with collective filtering.* These components correspond to a group of agents with the same interest or with the same taste in information selection. The joint filter can be considered as a *recommendation* of the group for its members to select from the information pieces.

The filters either can remain fixed during the functioning of the network or they can dynamically change. In the latter case the team members - depending on the information they have available at some moment of time - change the context conditions representing the input/output selector languages. Thus, at some moment the team of the agents can recommend new rules of information selecting, i.e. they *collaborate in determining filtering conditions.*

We should note that not only the filters but the teams can change during the functioning of the NLP systems: agents are allowed to migrate among the groups or they are allowed to join more than one team. This often takes place in real life: people modify their interest or, simply, to obtain some necessary information they join some new group of interest.

Variants of communication protocols lead to a wide variety of classifications of networks of language processors. If each rewriting step is followed by a communication step, then we speak of NLP systems with language processors communicating by command. If the rewriting at the component continues until a previously prescribed state (a state with a request for communication) is obtained and the communication step takes place afterwards, then we speak of networks of language processors communicating by request.

The above general model offers a language theoretic framework for modelling self-organizing, adaptive, evolving networks of (computational) agents. It is easy to see that it can be considered as a syntactic model of social networks with collaborative filtering of information or a syntactic approach to distributed /cooperative text processing systems realized on networks. The model, because of its general set-up, captures and can be extended to capture features of several extensively and intensively studied variants of networks: as a future example, the dynamic activity pattern of restricted features of the Internet could also be modelled in this way.

NLP systems are both computational and language identifying devices. Both their computational power and computational complexity and their language theoretic properties, including descriptive and size complexity, are of interest. (Languages can be associated with networks of language processors in various manners: for example, we distinguish a master component and take, as the corresponding language, any string that appears at this component during the computation.) In addition, since during their functioning NLP systems determine dynamically changing string multitudes, complexities concerning spatiotemporal dynamics of the emerging string collections are of particular interest. Studying, for example, the occurrence of waves or overloaded situations at the nodes, in the case of string multitudes of networks of language processors can lead to a deeper insight into the nature of distributed and parallel symbol processing and can help in understanding and modelling emerging phenomena in the case of networks like Internet.

3 A formal model

To illustrate the informal framework, we present the formal definition of a variant, called a *network of parallel language processors with teams with collective filtering* (a $TNLP\_F0L$ system, for short). The notion was formulated by some modifications of the notion of a network of parallel language processors ([6]).

In this case the language processors at the components are F0L systems, so-called interactionless Lindenmayer systems with a finite set of axioms, which are, roughly speaking, context-free grammars with a totally parallel way of derivation. (Originally, these systems were introduced for modelling developmental systems in terms of formal grammars, motivated by theoretical biology. The reader can find detailed information on Lindenmayer systems in [9].)

We assume that the reader is familiar with the basics of formal language theory. We list here only some notions which are necessary to follow the ideas of the formal contructions; for more details confer to [9].

For an alphabet V, V^+ denotes the set of all nonempty strings (words) over V. The empty string is denoted by λ, V^* stands for $V^+ \cup \{\lambda\}$. A language L is a subset of V^*.

An $F0L$ system is a triple $H = (V, P, F)$, where V is an alphabet, $F \subset V^*$, is a finite set of axioms, and P is a finite set of productions (rules) of the form $a \to v$, where $a \in V$ and $v \in V^*$. Moreover, production set P is complete: for every $a \in V$ there is a rule of the form $a \to v$, $v \in V^*$ in P. If F consists of exactly one string, then we speak of an $0L$ system. The direct derivation relation in an $F0L$ system $H = (V, P, F)$ is defined as follows: for $x, y \in V^*$ we write $x \Longrightarrow_P y$ if $x = a_1 \dots a_n$, $y = z_1 z_2 \dots z_n$, $a_i \in V$, $z_i \in V^*$, $1 \leq i \leq n$, and $a_i \to z_i \in P$.

In the following we define the *network of parallel language processors with teams with collective filtering.* We use some simplifications with respect to the general model: the number of the components, the rewriting rule sets of the language processors, the filters and the teams remain unchanged during the functioning of the system, the processors work in synchronized manner and the components check the strings to be communicated by using context conditions. Moreover, each agent is member of exactly one team.

Definition 3.1

A *network of parallel language processors with teams with collective filtering* of degree n, $n \geq 1$, (a $TNLP\_F0L$ system, for short) is a construct

$$, = (V, (\rho_1, \sigma_1, t_1), \dots, (\rho_n, \sigma_n, t_n), R),$$

where

- V is an alphabet (the alphabet of the system),
- ρ_i and σ_i, $1 \leq i \leq n$, are context conditions over V^* (computable mappings from V^* to $\{\underline{true}, \underline{false}\}$), called the *exit filter* and the *entrance filter* recommended by the i-th team to the members of the team, respectively,
- $t_i = (c_{i,1}, \dots, c_{i,r_i})$, $1 \leq i \leq n$, $r_i \geq 1$, called a team of components of the system (the i-th team), where
- $c_{i,j} = (P_{i,j}, F_{i,j})$, $1 \leq i \leq n$, $1 \leq j \leq r_i$, called a component of the network, the (i,j)-th component, where
- $P_{i,j}$ is a finite set of $F0L$ rules over V, the production set of the (i,j)-th component and
- $F_{i,j} \subset V^*$ is a finite set, the set of axioms of the (i,j)-th component, $1 \leq i \leq n$, $1 \leq j \leq r_i$,
- $R \subseteq \Pi \times \Pi$, where $\Pi = \{c_{1,1}, \dots, c_{1,r_1}, \dots, c_{n,1}, \dots, c_{n,r_n}\}$, called the neighbourhood relation of the components of , .

The components represent agents, which by using their sets of rewriting rules can update the textual information they have. Moreover, they form groups (teams), members of which have the same filtering conditions for selecting the information to be communicated and received.

The $TNLP$ system is functioning by changing its states.

By a *state* of a $TNLP\_F0L$ system , $= (V, t_1, \dots, t_n, R)$, $n \geq 1$, we mean a tuple $s = (L_{1,1}, \dots, L_{1,r_1}, \dots, L_{n,1}, \dots, L_{n,r_n})$, where $L_{i,j} \subseteq V^*$, $1 \leq i \leq n$, $1 \leq j \leq r_i$.

L_{i_j} is called the state of the (i,j)-th component and it represents the set of strings which are present at component (i,j) at that moment.

$s_0 = (F_{1,1}, \dots, F_{1,r_1}, \dots, F_{n,1}, \dots, F_{n,r_n})$ is said to be the *initial state* of the system.

A state can change either by a *rewriting step* or by a *communication step*. When a rewriting step takes place, then every component derives from each available string a new one, by applying its productions in the $F0L$ manner. Thus, in this case the number of strings available at the components does not change: each agent has the same number of strings as it had before the rewriting step.

At a communication step, each component (i,j) receives a copy of all strings that are present at some of its neighbourhood components, say, component (k,l) and are able to pass the exit filter of component (k,l) - this is the exit filter recommended to use by team k - and the entrance filter of component (i,j) - the entrance filter recommended by team i for receiving messages. (These strings satisfy context conditions ρ_k and σ_i).

Definition 3.2

Let , $= (V, t_1, \dots, t_n, R)$, $n \geq 1$, be a $TNLP\_F0L$ system.

Let $s_1 = (L_{1,1}, \dots, L_{1,r_1}, \dots, L_{n,1}, \dots, L_{n,r_n})$, and $s_2 = (L'_{1,1}, \dots, L'_{1,r_1}, \dots, L'_{n,1}, \dots, L'_{n,r_n})$ be two states of , . We say that

- s_1 directly changes for s_2 by a *rewriting step*, written as

$$(L_{1,1}, \dots, L_{1,r_1}, \dots, L_{n,1}, \dots, L_{n,r_n}) \Longrightarrow (L'_{1,1}, \dots, L'_{1,r_1}, \dots, L'_{n,1}, \dots, L'_{n,r_n})$$

  if $L'_{i,j}$ is the set of words obtained by performing a derivation step on each element of $L_{i,j}$ by production set $P_{i,j}$ in the $F0L$ manner, $1 \leq i \leq n, 1 \leq j \leq r_i$,

- s_1 directly changes for s_2 by a *communication step* in , , written as

$$(L_{1,1}, \dots, L_{1,r_1}, \dots, L_{n,1}, \dots, L_{n,r_n}) \vdash (L'_{1,1}, \dots, L'_{1,r_1}, \dots, L'_{n,1}, \dots, L'_{n,r_n})$$

  if for every i, $1 \leq i \leq n$, and j, $1 \leq j \leq r_i$,

  $L'_{i,j} = L_{i,j} \cup \{v \mid v \in L_{k,l},\ \rho_k(v) = \underline{true} \text{ and } \sigma_i(v) = \underline{true}, 1 \leq k \leq n, 1 \leq l \leq r_k, (k,l) \neq (i,j), (c_{i,j}, c_{k,l}) \in R\}$.

Notice that according to the above definition an agent in team i is allowed to receive messages from another agent of the same team.

A sequence of subsequent states determines a computation in , .

Let , $= (V, t_1, \dots, t_n, R)$, $n \geq 1$, be a $TNLP\_F0L$ system. By a *computation* C in , we mean a sequence of states $s_0, s_1, \dots$, where

- $s_i \Longrightarrow s_{i+1}$ if $i = 2j$, $j \geq 0$, and
- $s_i \vdash s_{i+1}$ if $i = 2j + 1$, $j \geq 0$.

Let , $= (V, t_1, \dots, t_n, R)$, be a $TNLP\_F0L$ system.

The *language* $L(,)$ determined by , is
$L(,) = \{w \in L_1^{(s)} \mid (F_{1,1}, \dots, F_{n,r_n}) = (L_{1,1}^{(0)}, \dots, L_{n,r_n}^{(0)}) \Longrightarrow (L_{1,1}^{(1)}, \dots, L_{n,r_n}^{(1)}) \vdash (L_{1,1}^{(2)}, \dots, L_{n,r_n}^{(2)}) \Longrightarrow \dots \Longrightarrow (L_{1,1}^{(s)}, \dots, L_{n,r_n}^{(s)}),\ s \geq 1\}$.

4 On the power of TNLP systems

Networks of language processors are language determining (computational) devices, therefore the question how large language classes (how complicated string communities) can be computed by their particular variants is one of the most important questions. Especially interesting are those NLP systems which are of considerable computational power and at the same time with extremely simple presentation. Simplicity in this case means, among other things, restricted size parameters of the network (a small number of components), poor power of the language theoretic operation represented by the language processor (restricted capabilities of the agents), homogenity of the components, simple communication protocol and simple (regular, subregular) filter languages.

Networks of language processors with $F0L$ systems as components and with filter languages defined by regular context conditions form a computational device equally powerful to the Turing machine ([6]). (To pass a regular filter, the string have to be an element of the regular language identifying the filter.) Morevover, it can be shown that in the case of regular filters a bounded number of components is sufficient to reach computational completeness. The same results can be derived for TNLP systems. Thus, TNLP systems with parallel language processors even with very simple presentation and with relatively simple filtering are able to process very complicated string collections.

Classes of languages determined by several kinds of networks of language processors based on different language theoretic operations have been studied in detail: it was shown, for example, that networks of language processors with regular filters and with context-free grammars as language processors or with language processors based on language theoretic operations simulating the recombinant behaviour of DNA strands or with operations corresponding to point mutations (splicing, cutting and recombination, insertion, deletion, replacement, etc.) provide universal computing devices. (For an overview on the area the interested reader is referred to [2].)

5 String collections of TNLP systems

Networks of language processors are devices not only for describing the dynamics of languages at the components but they also provide tools for characterizing multisets of strings. Properties of these string collections are of particular importance in those cases when not only the information piece itself (for example, the arriving message), but the number of its available copies is of interest. Since the notions related to these networks of string multiset processors (NMP systems) are isomorphic to the notions concerning NLP systems, we omit the explicit definitions. We only note that in this case the computing devices located at the components operate on such collections of strings where the strings are allowed to have multiple (a finite number of) occurrences of the same copy.

In [6] it was shown that the growth of the number of strings present during the computation at the components of an NMP system which has random context filters and deterministic F0L systems as components can be described by the growth function of a D0L system. (A D0L system is an 0L system with exactly one production for each letter a of the alphabet at the left-hand side. A random context filter checks the string according to the presence/absence of some symbols. The growth function of a D0L system orders to each natural number n the length of the word generated by the system at the n-th step of the derivation.)

The proof is based on the following simple considerations: since $D0L$ systems define homomorphisms, therefore if we know how many strings with a fixed alphabet are present at some component, then we are able to give the number of strings with the same alphabet obtained after performing a rewriting step at the component. Moreover, because at communication steps we check the presence/absence of some symbols in the strings, we are able to decide whether a string with a fixed alphabet can pass a filter or not. Thus, at any state of the computation we can represent the multiset of strings at some component by the multiset of their alphabets, and, we can construct a $D0L$ system such that the multiset of the letters of the word of the $D0L$ system at step t is equal to the multiset of the alphabets of the strings present at some component (at the components) at a corresponding step of computation in the network.

Using this proof technique, the same result can be given in the case of TNMP systems with deterministic F0L components and random context filters. Moreover, we note that not only the growth of the whole string community, but also the growth of the number of strings at the individual components and teams can be calculated. By the theory of D0L growth functions we can derive several interesting properties of the emerging string collections at the TNMP systems. We know, for example, that the growth of the string population (at some team or at some individual component) is either polynomially bounded or exponential and this is a decidable property.

6 Final remarks

In this paper we briefly discussed a general framework which provides language theoretic approach for describing the behaviour of agents and agent communities which use networks for cooperation and communication. We hope that the theoretical model can help in developing tools for designing languages supporting collaborative text processing via networks.

References

[1] Communication of the ACM, March 1997, Vol. 40., No. 3. Special issue: Recommender Systems. Linking users by similar interest.

[2] E. Csuhaj-Varjú, Networks of Language Processors. EATCS Bulletin, 63 (1997), 120-134.

[3] E. Csuhaj-Varjú, J. Dassow, J. Kelemen and Gh. Păun, Grammar Systems: a Grammatical Approach to Distribution and Cooperation. Gordon and Breach Science Publisher, London, 1994.

[4] E. Csuhaj-Varjú, L. Kari and Gh. Păun, Test Tube Distributed Systems Based on Splicing. Computers and Artif. Intelligence 15(2-3) (1996), 211-232.

[5] E. Csuhaj-Varjú, J. Kelemen and Gh. Păun, Grammar Systems with WAVE-like Communication. Computers and Artif. Intelligence 15 (5) (1996), 419-436.

[6] E. Csuhaj-Varjú and A. Salomaa, Networks of Parallel Language Processors. In: New Trends in Formal Languages. Control, Cooperation and Combinatorics, (Gh. Păun, A. Salomaa, eds.), LNCS 1218, Springer Verlag, Berlin-Heidelberg-New York, 1997, 299-318.

[7] L. Errico and C. Jesshope, Towards a new architecture for symbolic processing. In: Proc. Conf. Artificial Intelligence and Information-Control Systems of Robots' 94, (I. Plander, ed.), World Scientific, Singapore, 1994, 31-40.

[8] S. E. Fahlman, G. E. Hinton, T. J. Seijnowski, Massively parallel architectures for AI: NETL, THISTLE and Boltzmann machines. In: Proc. AAAI- Natl. Conf. on AI., William Kaufman, Los Altos, 1983, 109-113.

[9] Handbook of Formal Languages. Vol. I-II-III. (G. Rozenberg, A. Salomaa, eds.), Springer Verlag, Berlin-Heidelberg-New York, 1997.

[10] C. Hewitt, Viewing Control Structures as Patterns of Passing Messages. J. of Artificial Intelligence 8 (1977), 323-364.

[11] W. D. Hillis, The Connection Machine, MIT Press, Cambridge, 1985.

The TREVI Project

Personalized Information Filtering, Linking, and Delivery for the News Domain

Reginald Ferber and Costas Tzeras, GMD - IPSI, Dolivostr. 15,
D-63293 Darmstadt, Germany, {ferber,tzeras}@darmstadt.gmd.de,
http://www.darmstadt.gmd.de/~ferber http://www.darmstadt.gmd.de/~tzeras

The goal of the TREVI project (Text Retrieval and Enrichment for Vital Information) is to offer a solution to the problem of "information overflow", i.e. the problem experienced by companies and individuals in extracting useful information from distributed textual information sources. These information sources are available through public distribution channels such as the Internet and the World Wide Web, or through proprietary networks. At the same time, more and more archival or encyclopedic data collections are becoming available in electronic format, providing background knowledge to particular business domains.

The TREVI approach aims to filter information from streams of incoming news (information sources) based on individual user profiles. Furthermore, TREVI aims to enhance the filtered information by enrichment with background data sources in accordance with user profiles. The filtered and linked information will be presented in a coherent and comprehensible way to end-users (document publication).

TREVI is an ESPRIT joint project (ESPRIT Programme 23311) of GMD-IPSI with

- Economisch Instituut Tilburg (EIT), Netherlands;
- FEND Association, Spain;
- ITACA s.r.l., Italy;
- Lyras Shipping LTD, United Kingdom;
- REUTERS LTD, United Kingdom;
- SARENET SA, Spain;
- Vrije Universiteit Brussel (VUB), Brussels.

It will run from January 1997 to June 1999.

GMD-IPSI's work on TREVI is divided into two parts: (1) Text Enrichment and (2) Document Publication. GMD-IPSI also assists the project partners in specifying a representation formalism for user profiles.

TREVI will be applied to four test environments:

- The Italian Health online service (ARAKNE) that provides news, research results, and information documents for the medical domain from various sites. As background material there will be archives of this service.
- The ECO PRENSA service that provides Spanish abstracts of newspaper articles from the economics domain. As background material there will be databases with information from the stockmarket.
- An experimental subset of Reuters news service. As background material there will be a set of historical information and selected news articles, stockmarket and company information.
- The distribution of business circulars within the Lyras shipping company. These circulars include news, guidelines, and business informations that have to be directed to the appropriate persons within the company.

The incoming information streams will be heterogeneous. They will rank from unstructured texts to news that are structured by different fields like author, city of origin, subject etc. The background material will also be heterogeneous. It will include unstructured text documents, weekly structured material and databases as highly structured information.

The main tools for filtering are a lexicon system and the user profiles. The lexicon system is a kind of enriched ontology based on WordNet that allows to specify concepts. WordNet is enriched with specific information from the domain and with linguistic and terminology information for parsing and tagging.

The user profiles contain various types of information: Content information is specified by concepts from the lexicon. Some metadata specify formal information like sources to be used, cost limits, time and geographical restrictions as far as they can be identified in structured sources. A second part of the user profiles is information concerning strategies of enrichment. This can be the time at which the enrichment shall be made, different search strategies, and formal restrictions or properties for background material like source, length, price, age... The last package of information concerns the selection and configuration of modules used for the specific user. This selection depends on the information sources, the availability of specified lexical information and the retrieval methods to be used. It will affect the speed and the costs of linking. Probably there will be a fixed selection of configurations for the most likely scenarios.

The user profiles will be created either by information experts or brokers for their clients, or by expert clients themselves. Such experts will be able to change their profiles temporarily or permanently. There will be also some predefined profiles for casual users.

The publication and user interaction component of TREVI will supply three different modes to be selected depending on the user habits and the network and hardware situation. In e-mail mode the filtered news and the respective background information will be send to the user in fixed time intervals or upon a arrival. In the two other modes - the HTML and APALO mode - the filtered news will be collected and shown when the user logs in. In this case users can select if they want to see all items that arrived since they logged in last or only those from a given time period. The APALO mode is named after a layout system developed by GMD IPSI that puts strong emphasis on a structured and content sensitive presentation of text and images. Both the HTML and the APALO mode will provide personal archives, to store and retrieve documents. Retrieved documents can be used to select similar information from the news stream or the background material. Advanced users can select retrieval strategies. In addition they can use relevance feedback and a profile editor to change their user profiles.

Using of multiple data source for information filtering: first approaches in the MedExplore project.

Emmanuel Nauer, Jacques Ducloy, Jean-Charles Lamirel

CRIN-CNRSet INRIA-Lorraine
Bâtiment LORIA

615, rue du Jardin Botanique

B.P. 239

F-54506 Vandoeuvre-les-Nancy Cedex

E-mail : {Jacques.Ducloy, Emmanuel.Nauer, Jean-Charles.Lamirel}@loria.fr

Keywords

Internet, Information Retrieval System, navigation graph, information statistical analysis, single link clustering, structured data use, multiple source.

Introduction

One of the major challenge of modern information retrieval and of technological survey is the mastering of the use of multiple and various data sources.

In this paper, we first describe the experimental workbench we have set up in the framework of the MedExplore project. This workbench allows both to merge and to cross data issued from multiple funds, including the Internet.

We will detail, an original navigational graph building method based on structured data. This method provides the user with a hypertextual access through thematics ordered by different generality levels. These thematics play the role of guidelines to help the user to formulate a query on the net, whenever his competence level or his type of need in the investigation field.

We will finally give various other samples from the MedExplore project that illustrate the usefullness of the crossing of strutured data. In that part, we will also describe our first heuristics to achieve contextual search on different survey areas.

1. The MedExplore Project

1.1. General overview

The aim of MedExplore is to give a group of experts confronted with an unforeseen field the mastery of its terminological resources together with a synthetic and deeper knowledge of the state of the art of that latter. This task will be achieved by creationing of a system of investigation.

Such a system (see below, figure 1) allows a user to navigate through concept graphs and to manipulate conjointly various pieces of information (large international databases, local source documents, raw information from the INTERNET), written in different languages.

We have chosen to begin our experimentations on biomedical fields because of the large amount of what we call "structuring" funds. Indeed, such funds like MEDLINE, EMBASE or PASCAL possess a homogeneous indexing and also a "quasi" knowledge based representation when they are associated with projects like UMLS.

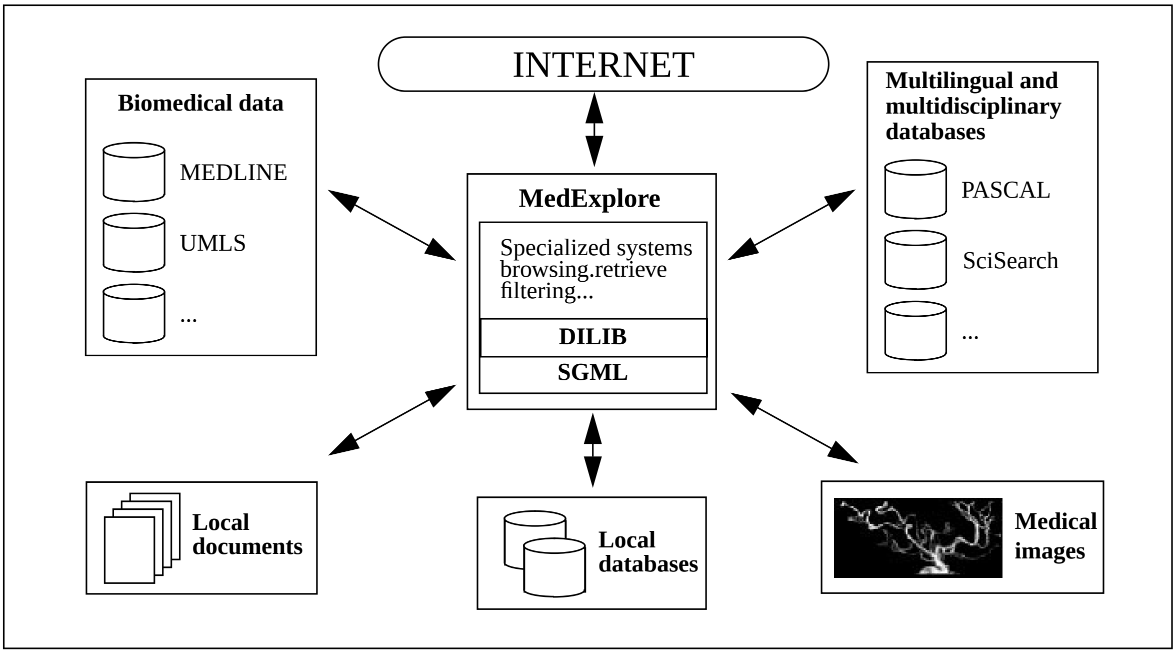


Figure 1 : Overview of the MedExplore project

1.2. Using MedExplore : basic principles

As we have mentioned before, we mainly aim at providing access for different types of users (from the traditional "end-user" to the "expert in data analysis") to a server which deals with different data sources in a coherent and homogeneous way (see figure 2).

For that, we have to solve the different levels of inter-operability between heterogeneous data. SGML/XML brings us a good answer for the "codification - structuration" level. Now we have to deal with more semantic levels. As we work in specialised area, we have chosen to simplify this problem by defining a core vocabulary which contains a limited number of terms (between 100 and 300) and which represents a kind of semantic gateway between the different databases.

Such a lexicon can be generated by automatic tools (for instance clusterisation) and improved by a specialist if necessary.

1.3. First approaches for the use of structured data for information filtering

The use and the crossing of structured data from multiple sources tend to facilitate IR in several ways. We will describe hereafter more precisely some of the main advantages of this approach which was dem-

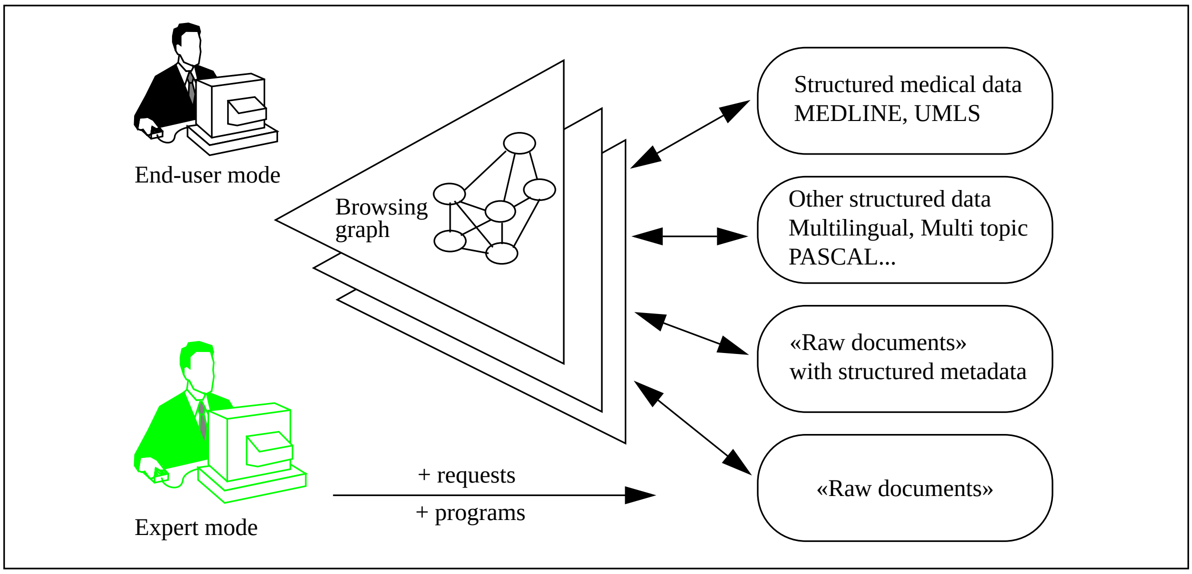

End-user mode
Browsing graph
Expert mode
+ requests
+ programs
Structured medical data MEDLINE, UMLS
Other structured data Multilingual, Multi topic PASCAL...
«Raw documents» with structured metadata
«Raw documents»

Figure 2 : Homogeneous access to the data through a navigation graph

onstrated in our MedExplore experimentations.

Thematic access to WWW

In [NAU97], we have already described how simple bibliographic references issued from Medline allow the automatic building of specialised investigation system. The proposed system was composed of a thematic navigation graph which represents the interface between the user and WWW. The main task of the system was to choose automatically the field vocabulary to lead the user to overcome his limits in terms of vocabulary, knowledge and memory. The user could therefore simply formulate the queries to submit to the search engines through terms selection. This conceptual view of query formulation has been also adopted by AltaVista [ALT97] with the "Refine" options (formerly "Live Topics"), or by Excite [EXC97], which proposed a word set to extend the initial query. Nevertheless, our approach seems more accurate, because it works on field knowledge instead of statistical links. Indeed, it has turned out that these last links are often inappropriate (verbs, idioms,...)

Multilingual access

The complementary use of UMLS allows us to provide the system with multilingual capabilities, allowing then thematic access in different languages. Therefore, the previously described information gateway, trough keywords, also plays the role of finding linguistic equivalences for the selected terms considering the interrogation language. For example, a French user could access to a French thematic graph whilst interrogating transparently the search engine in English. This is a convenient solution for an unskilled user with the English translation of his native query terms.

Complementary information on documents

Remote documents found during query sessions could be dynamically enriched through hypertext inverted links to the graph themes whose role represents complementary information for the user. As a matter of fact, they allow him both to operate accurate thematic linking between the Internet documents and the themes constituting the navigation graph and to retrieve directly bibliographical information which could be useful for his explanation(s).

Generating Dublin Core Metadata.

From the computer point-of-view, the data-crossing process produces a set of resources, such as tables, which can be also used in a generation process.

We are working in that way for the generation of a server allowing browsing through a collection of Medical Images. Each image is described in French with a short set of information :

```
<doc>
   <id>lethor/007_001</id>
   <auteur><e>Lethor JP</e><auteur>
   <specialite>Cardiologie infantile</specialite>
   <tech><e>Radiographie</e><tech>
   <organe><e>Coeur</e><e>Poumon</e></organe>
   <patho><e>Tétralogie de Fallot</e></patho>
   <motif>rupture de patch infundibulaire</motif>
   <age>12 ans</age>
</doc>
```

For this image, we can generate an HTML page whose metadata are the following:

```
<meta  name="DC.title" lang="fr" content="Image : tétralogie de Fallot">
<meta  name="DC.creator" content="Lethor JP">
<meta  name="DC.subject" lang="fr"
       content="Coeur, COEUR (RADIOGRAPHIE ), Poumon, POUMON (RADIOGRAPHIE ),
       Radiographie, Tétralogie de Fallot">
<meta  name="DC.subject"
       content="Heart, Heart_Radiography, Lung, Lung_Radiography, Radiography, Tetralogy of Fallot">
<meta  name="DC.subject" scheme="MESH"
       content=" Heart, Lung, Radiography, Tetralogy of Fallot">
```

Help for technological and scientific survey

The mastering of the analysis of great information resources available on Internet has been, for the last recent years, one the main challenge of technological and scientific survey. The information being on the net, thanks to constant evoluation, allows recent data analysis, in opposition with studies based on classical documentary databases [AND97][ROS93].

Besides, search engines on the net (AltaVista, Excite, Lycos, etc.) suffer from the same defects as classical documentary systems [DUB94] : the absence of a reliant indexing of their documents cause them difficulties to give back relevant documents to the user. In fact, these systems propose most of the times as a query response : - too many documents ;
- documents with low relevance considering the user’s real need.

Therefore, achieving a successful Internet search obliges introducing search mechanisms using knowledge which is specific to the interrogation field. This mechanism, we have call "Contextual Information Filtering", has also been implemented in the framework of the project MedExplore. It is described hereafter (see ???).

2. 2. A technical base for MedExplore : the DILIB workbench [DIL97]

The engineering techniques used for MedExplore is based on the generalisation of SGML codification allowing the use of SGML toolboxes and linguistic modules libraries.

Therefore, we have developed DILIB, an SGML workbench [DUC94] which contains a set of basic components to build Information Retrieval Systems

2.1. SGML, homogenisation of information

We use to convert all information in an SGML markup [ISO86] whose structure is very close to the original one. For instance a downloaded record issued from MEDLINE such as:

```
AN : 96081277
TI : Orthotopic pulmonary valve replacement with a homograft.
AU : Saha K,Iyer KS, Sharma R, Bhan A, Airan B, Venugopal P
CS : Department of Cardiothoracic and Vascular Surgery, All India Institute of Medical Science
JN : J Heart Valve Dis CP : (ENGLAND)
PY : Mar 1995
VO : 4 (2) p187-91
...
```

becomes:

```
<MEDLINE>
    <AN>96081277</AN>
    <TI>Orthotopic pulmonary valve replacement with a homograft.</TI>
    <AU><e>Saha K</e><e>Iyer KS</e><e>Sharma R</e><e>Bhan A</e><e>Airan B</e><e>Venugopal P</e>
    </AU>
    <CS>Department of Cardiothoracic and Vascular Surgery, All India Institute of Medical Sciences</CS>
    <JN>J Heart Valve Dis</JN>
    <CP>(ENGLAND)</CP>
    <PY>Mar 1995</PY>
    <VO>4 (2) p187-91</VO>
    ...
</MEDLINE>
```

In some case, it may be interesting to carry out little transformation on some data sets. For instance, we need to handle multilingual information coming from UMLS whose records look like that :

```
C0017379|ENG|P|L0017379|PF|S0022690|Carriers, Genetic|
C0017379|ENG|P|L0017379|VW|S0044411|Genetic Carriers|
C0017379|ENG|P|L0017379|VWS|S0022684|Carrier, Genetic|
C0017379|ENG|P|L0017379|VWS|S0044407|Genetic Carrier|
C0017379|POR|P|L0436728|PF|S0561010|TRANSPORTADORES GENETICOS|
C0017379|SPA|P|L0447330|PF|S0571612|PORTADORES GENETICOS
```

where C0017379 identifies a unique concept with an English prefered form (*Carriers, Genetic*), various usual forms and some translations.

For an easier data management, it is more convenient to group all information related to a particular concept, which is originally distributed in a table format, into one SGML record like that :

```
<CONCEPT>
    <CUI>C0017379</CUI>
    <TP><PF>Carriers, Genetic</PF>
           <VW>Genetic Carriers</VW>
           <VWS>Carrier, Genetic</VWS>
           <VWS>Genetic Carrier</VWS>
    </TP>
    <VL l="POR">
    <TP><PF>TRANSPORTADORES GENETICOS</PF></TP>
    </VL>
    <VL l="SPA"><TP><PF>PORTADORES GENETICOS</PF>
    </TP></VL>
</CONCEPT>
```

In the same way, we apply such transformation on all the managed data, obtaining an SGML mark-up on which we can apply the associated engineering possibilities.

2.2. Handling SGML information with DILIB

DILIB provides a set of tools to handle SGML or XML elements. They are available at different programming levels.

For instance, if you want to add the key-word "*AIDS*" as an element tagged with <e> to an SGML element which is pointed by "kw" variable in a C program, you have to write:

```
SgmlAddChild (kw, SgmlCreateLeaf("e", AIDS));
```

In the same way, you can use shell commands to handle sets of records. We have introduced a "path pattern mechanism" to specify a set of elements into a given document. For instance, if you want to select records which contain "*AIDS*" as a subpart of keywords and print the corresponding titles, you just have to write:

```
SgmlSelect -g MEDLINE/KW/e#AIDS -g MEDLINE/TI -p @g2
```

(where "-g" is used by analogy with *grep* and *@g2* identifies the 2nd "g" sub-command)

As HTML and Dublin Core deals with SGML, it becomes very easy to generate metadata in a programming environment. For instance, the following program :

```
SgmlNode *meta;
meta= SgmlCreateEmptyMark("META");
SgmlSetAtt(meta, "name", "DC.subject");
SgmlSetAtt(meta, "content", "AIDS");
```

will produce :

```
<META name="DC.subject" content="AIDS">
```

2.3. Information analysis with MedExplore/DILIB

Now, if we want to know the vocabulary which will be appropriate to retrieve relevant documents or to produce significant metadata contents, we have to analyse a set of information.

For that purpose, DILIB also contains a set of basic components to build customised information retrieval systems. These tools allow a global analysis of large sets of information. Up to now, our tools have been mainly based on basic statistical approaches. An illustration of such an approach is the navigation graph building mechanism (clusterisation) described hereafter.

Clusterisation

The navigation graph building is based on a single link clustering method [JAR71]. This method works by iterating on keywords associations issued from the documents, ordered by decreasing relevance. Similarly to Michelet [MIC88], we choose the equivalence coefficient as the statistical indicator to order the keywords associations because it weights the associations importance relatively to their 2 component terms.

This coefficient is given by the following expression :

$$a_{ij} = \frac{f_{ij}^2}{f_i f_j} \qquad (1)$$

f_{ij} being the cooccurence count of keywords i and j in the documents and, f_i and f_j being respectively the i and j keywords occurence counts in the same documents.

However while making experiments, we found that the direct use of the equivalence coefficient on the whole association set, ordered by pertinence [MIC88], almost always inhibits the characterisation of the most general themes of a domain. That phenomenon mainly occurs when the field is strongly structured and composed by numerous very specific and very coherent subfields. Moreover, in some intermediate situations where the subdomains structuration is less strong, the direct use of said equivalence coefficient may lead to an uncontrolled mixture of general and specific themes.

To cope with the above described problems we propose an original two-step clustering method (see figure 3) :

The first step consists in establishing the generality level of the clusterisation by selecting the core associations set of the clusterisation thanks to the keywords cooccurrence count. An associations set with high cooccurrence count will always lead to general themes. Conversely, an associations set with low cooccurrence count will always lead to specific themes.

The second step consists in applying the classical single link clustering algorithm on the selected associations, reordering them by equivalence coefficient.

Finally, this method gives us the opportunity to adapt the navigation graph building to different types of users and to their different needs. The general graph will then be dedicated to the novice users with, most often, no specific knowledge about the explored field, limited vocabulary and general information needs. Conversely, specific graphs will be devoted to field specialists with more technical vocabulary, more elaborated knowledge and more focused needs.

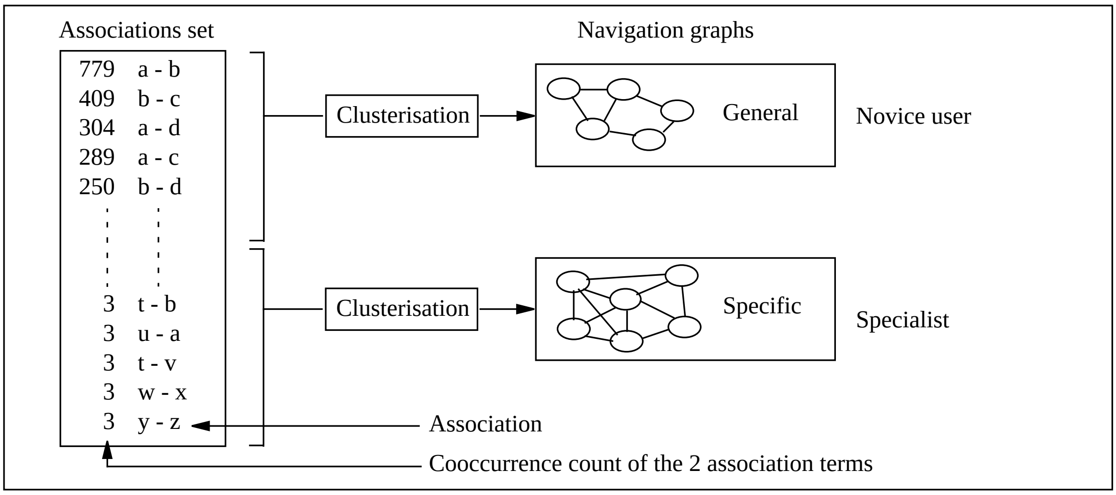


Figure 3 : Original two-step single link clustering

To order the resulting themes of each graph thanks to their generality level, we use a generality indicator which could be described by the below basic formula :

$$g_C = \frac{1}{\mathrm{Card}(C)} \sum_{(i,\,j) \in C} f_{ij} \qquad (2)$$

C being a given themes of the graph and *Card(C)* being the number of associations of the theme *C*.

3. Improvement with MedExplore : Contextual Information Filtering

3.1. General overview of the system

Making an Internet search successful implies solving a typical problem of the IRS : maximizing the number of relevant retrieved documents (recall) whilst minimizing the number of non relevant retrieved documents (noise).

Therefore, we have choosen to use structured knowledge to established a context around the central subject of the research. This technic we have called "Contextual Information Filtering" will limit the response scope and favour the emergence of relevant data.

Our first experimentations in that domain led us to set up a three component system (see figure 4):

The goal of the first component is to detemine the context around the focus of interest : this step is strongly guided by the expected results. Nevertheless, the general principle both uses structured data (sometimes linked with a core vocabulary), and information analysis methods which are for now statistical ones. In a short future, we aim at integrating linguistic methods such as : extraction of nominal groups, thesaurus use, ...

The second component implements the interrogation of the search engines, through multiple queries. At this step, our very first attempts originally use a single query. Unfortunately, our numerous experimentations show that it was very difficult, if not impossible, to find the optimal query. In most cases, adding terms will precise the query and so discriminate relevant documents from the non relevant ones. Nevertheless, the introduction of a general word (which indexes a very large scale of documents on the net) will sometimes alter the results given by the search engines. Indeed, these general words will distort the discriminant context, causing the emergence of non relevant documents. Therefore we choose to submit several well-formed queries and to add a synthesis step in order to suppress the side effect of the awkward "bad" queries.

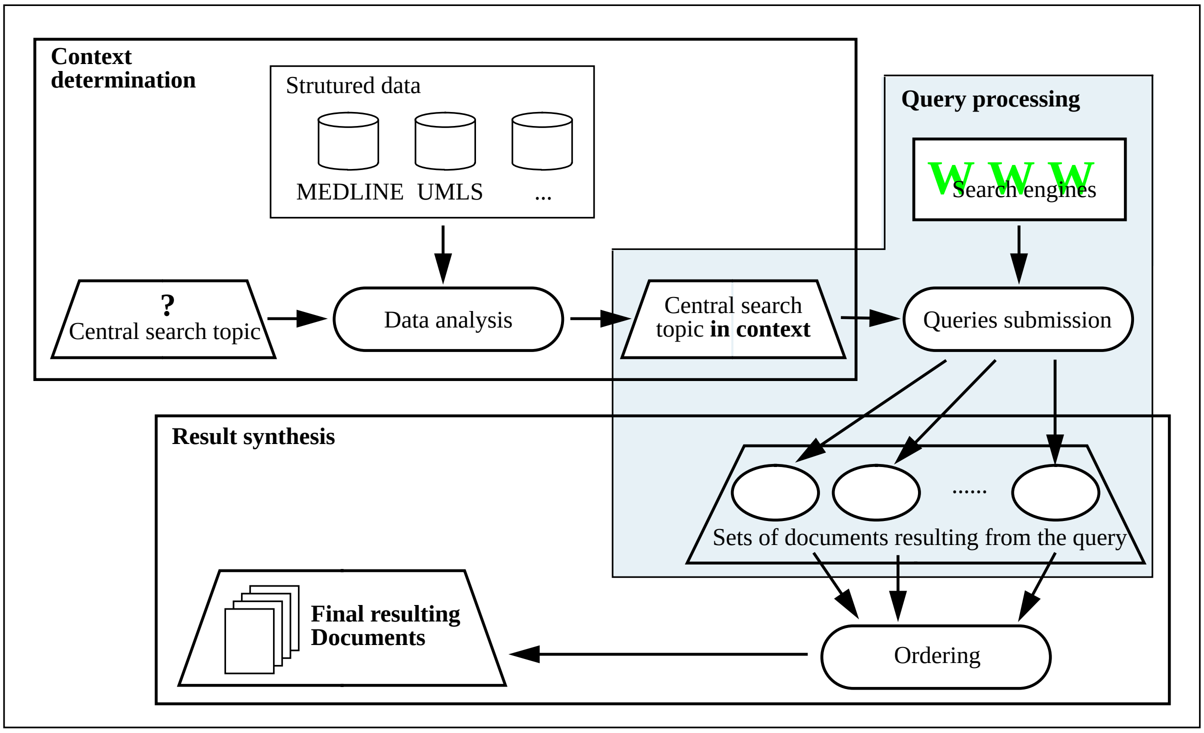


Figure 4 : Contextual Information Filtering System architecture

The last component deals with the results set to provide the user with a list of relevant documents. For that purpose, the weighted mean rank formula proposed in [LAM95] has prooven to be accurate. This ranking formula gives the weighted mean rank $r(\Re, d)$ of a document *d* thanks to a set of queries $\Re$ as :

$$r(\Re, d) = \frac{\sum_{i \in \Re} \alpha_i r(i, d)}{\sum_{i \in \Re} \alpha_i} \qquad (3)$$

where α_i is the occurrence weight of the document *d* in the query i and $r(i, d)$ is a function depending on the rank of the document *d* for the query *i*. When you want to eliminate from the final result the documents which have not been found relevant for one of the queries, the rank function could be set up to give an infinite rank to these documents. Conversely, to keep these documents, the rank function could be set up to attribute them the rank which is the nearest one of the one of the last documents found for the said queries.

We could mention that a single query might be considered sufficient in the very first step of a search session. In this case, the second component of our system will submit only one query to a search engine, and the ranking of the document will be directly given by the search engine query result (the third component will then be useless).

3.2. Samples of contextual information filtering

Large search on a single general keyword

For such an extraction of documents from the INTERNET, the principle consists in associating to the searched term a histogram of the most frequent terms cooccurring with this latter in the MEDLINE records. In databases, which are not indexed with keywords, we can also use full-text words, coming from the abstracts.

For instance, for the keyword "Newborn, Infant", in a local base dealing with "cardiology", the associated histogram (resulting from the abstract) will look like that :

> [59] patient - [42] pulmonary - [37] tetralogy - [33] fallot - [29] infant - [27] heart - [25] artery - [24] defect

A single query on a search engine, using this set of words, will then give us the expected results.

Research work of an author

The same kind of technics can be used to search for the work of an author. For instance, the vocabulary associated to Pr. JP Lethor from his bibliography on MEDLINE :

> [35] ventricular - [31] volume - [27] left - [26] dimensional - [20] coronary - [19] three - [16] method - [15] patient
> [13] defect - [13] image - [12] excised - [11] doppler - [10] artery - [10] echocardiography - [10] tau - [9] pressure

A single query using the author name complemented by this set of words will then give us, again, the required results.

Research area of an author

To generalize the search around a research thematic, a histogram could be obtained thanks to authors working on that thematic or/and by using documents relative to this latter. Unlike the previous cases where the histogram was used to overdefine the initial query, we only made use of the generated histogram without the elements (keyword or author) from which it was built.

This approach seems to be very appropriate to achieve technological survey where one may not always be faced with precise and non-evolutive area.

Conclusion

In the very first steps of the MedExplore project, we have hightlit the usefullness of crossing structured data, coming from multiple sources, by testing and implementing several functionalities of an integrated IRS. This functionnalities include the thematic and multilingual access to WWW, the complementary information about retrieved documents and the generation of metadata for standard web documents.

We now explore in a deeper way and try to generalise one of these functionnalities, which we have called "Contextual Information Filtering". We will attempt to improve our retrieval performances, through different ways :

- the study of different strategies for the "queries submission" part of our Contextual Information Filtering System will help us to determine how to submit a minimal number of queries. We will particularly focused our work on the context definition. For that, we will introduce technics coming from linguistic or even from statistics. We will also use other types of structured data like the ones coming from a thesaurus.

- the experimental determination of the best ordering function will allow us to maximize the recall and to minimize the noise.

- the analysis of the Internet documents with very close technics will lead us to determinate the accurate metadata (from the ones being present in the documents), which could then be used to set up a better context. This context being more suitable to the search engines will then improve the search performance on the net.

Bibliography

[ALT97] AltaVista. Information and search engine access : http://altavista.digital.com/

[AND97] Pascal Andrei. *Elaboration et traitement d’information complexe pour l’aide à la décision stratégique.* Phd Thesis in Information Scientifique et Technique, University of Marne-la-Vallée, 1997.

[DIL97] Information about the DILIB workbench : http://www.loria.fr/DILIB

[DUB94] Jacques Emile Dubois et Belhadri Messabih. Internet-web and ST data management : harmonization and new horizons. In *The Information Revolution : Impact on Science and Technology*, pp 43-56, 14th International ConferenceCODATA, 18-22 septembre 1994, Chambéry, France.

[DUC94] Jacques Ducloy, Jean-Charles Lamirel et Emmanuel Nauer. A workbench for bibliographical or factual data handling. In *The Information Revolution : Impact on Science and Technology*, pp 63-70, 14ème conférence internationale CODATA, 18-22 septembre 1994, Chambéry, France.

[EXC97] Excite. Information and search engine access : http://www.excite.com/

[ISO86] ISO 8879. Standard Generalized Markup Language (SGML), 1986.

[JAR71] N. Jardine et R. Sibson. *Mathematical Taxonomy*. Wiley, Londres et New York, 1971.

[LAM95] Jean-Charles Lamirel. *Applications d'une approche symbolico-connectionniste pour la conception d'un système documentaire hautement interactif : le prototype NOMAD.* PhD Thesis in Computer Science, University of Nancy I, 1995.

[MIC88] Bertrand Michelet. *L'analyse des associations*. PhD Thesis in Information Sciences, University of Paris 7, 1988.

[NAU97] Emmanuel Nauer, Jean-Charles Lamirel. *Environnement d'investigation sur WWW assistance à l'utilisateur par des connaissances fédérées*. In H2PTM'97, pp 101-113, 4th International Conference : Hypertexts and Hypermedia - Products, Tools and Methods, Hermes, September 1997.

[ROS93] Hervé Rostaing. *Veille Technologique et Bibliométrie : Concepts, Outils, Applications*.PhD Thesis in Sciences de l'Information et de la Communication, 1997.

The end of symbolic immortality: a non-monetarian collaborative cooperation model in an Internet based groupware service

Roland ALTON-SCHEIDL
Gernot TSCHERTEU
Austrian Academy of Sciences' Research Unit for Socio-Economics
(e-mail: firstname.lastname@oeaw.ac.at)

Abstract

Based on user requirements investigations of typical groupware users, we have elaborated a rating model for collaborative message filtering. In this model, evaluating contributions has a direct effect to the organisational structure of virtual co-operative groups. The model fosters self-organisation and vitalization. The dynamic incentive mechanism mirrors real group dynamics to virtual communities. In that sense, symbolic immortality of human beings (Baudrillard) is reduced for virtual activities, as symbolic capital has to be steadily renewed.

Introduction

A user requirements phase of the project Web4Groups, which was carried out by a consortium of European R&D and Industrial partners and has been funded by the European Commission within the "Telematics Application Programme", included archetypical user groups from the research and administration sector. The lack of quality of information in threaded discussion groups has been stressed by the representative of the user groups as an important issue. 43% of the users see collaborative rating as a reasonable way for better filtering and coping with the huge amount of messages in mailing lists or newsgroups.

Within the follow-up project SELECT [8] we propose an architecture for filtering mechanisms both for the world of messaging and documents. The model described in this paper is a candidate for implementation using the general purpose groupware tool Web4Groups.

Role Handling in Groupware

Groupware implementations offer a wide variety of handling roles. A complex role handling mechanism is not necessarily a guarantee for acceptance of a groupware system by the users. Notably, very simple role models are today used for group communication widely and successful: WWW and e-mail.

The authors describe a system of flexible and context-sensitive role assignment in the paper `Social Functions in Virtual Communities' [1]. Like in `real life' roles should change according to the social settings in which a person participates. It is not self-evident that usage rights should be assigned without regard to the social setting in which the usage takes place. CMC should be a socially differentiated cultural space, like cultural spheres in `normal life', and this would for example mean that a participant in a discussion should have certain rights, different from somebody participating in a market situation. Such a change from one social setting to another is described in Figure 1.

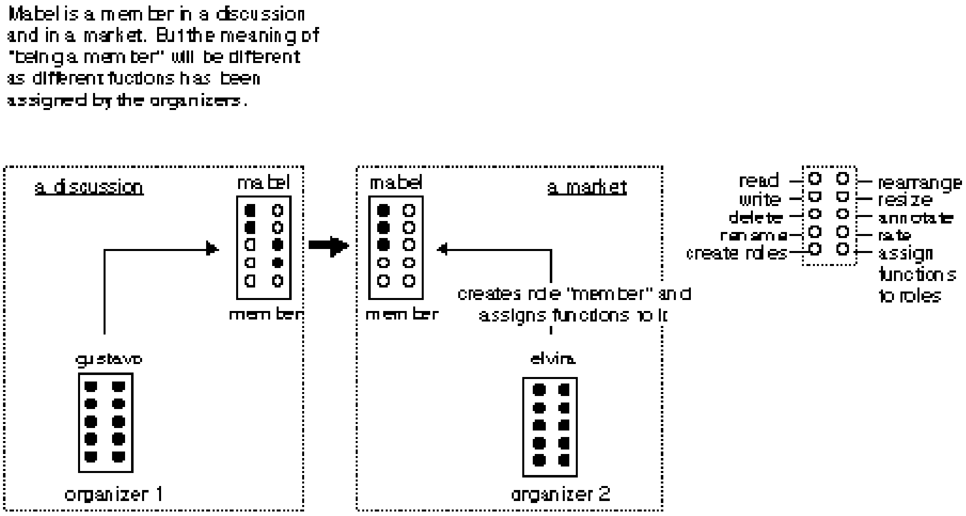

Mabel is a member in a discussion
and in a market. But the meaning of
"being a member" will be different
as different fuctions has been
assigned by the organizers.
a discussion
mabel
member
gustavo
organizer 1
mabel
a market
member
creates role "member" and
assigns functions to it
elvira
organizer 2
read
write
delete
rename
create roles
rearrange
resize
annotate
rate
assign
functions
to roles

Figure 1: Roles change in different virtual social settings

The point is that there are not only differences between roles but also differences between similar roles in different settings. Being a family member has different implications from being a member in the Houses of Parliament. This is self-evident in `normal life' and should also be the case in CMC.

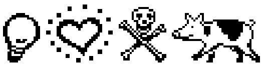
The idea: linking role assignment and rating

The next observation was that functions and rights should not only be assigned by organisers. In Web4Groups it was tried to avoid such an omnipotent role. The question arose whether rating could be a suitable way to distribute and obtain functional elements. The basic idea was: if a person participated very actively and wrote interesting contributions, s/he would get good ratings. Those could be collected in a personal account. When a person had collected a certain amount, s/he should be assigned new functional elements or a completely new role. Furthermore, persons would be able to gather "symbolic capital" and show it proudly to other users.

Linking ratings with functional elements and roles

Ratings may refer to different aspects of contributions. Therefore different ways of rating units may be used: e.g. a rating unit for innovative ideas: `light-bulbs'. If a contribution gets three `bulbs' it is supposed to contain a lot of brilliant, new thoughts.

Another rating unit - e.g. hearts - could be awarded for sympathetic contributions, a third one - maybe medals - for social competence, and so on.



After having for example received five bulbs, the participant should obtain the right to make annotations to other contributions, or, after having received ten `hearts', he or she would become `man or woman of the month', and find his or her face on the welcome page of that service. There are no limits to inventing new rating units and things that could happen after having received a certain amount of them.

In general one could distinguish between two forms of units depending on the consequences of good ratings:

<u>First</u> there could be some <u>automatic relation</u>:

E.g. after having received ten light bulbs somebody would automatically get the right to make annotations to contributions of other members.

<u>Second</u> there could be <u>free choice</u> among a set of additional rights:

be paid with three `pigs'. What would s/he buy with that `money'? There must be other services in the community that would be interesting for him/her, like the advice of another competent person, e.g. a tax consultant, a translator, an editor, and so on. Such a system of exchange makes sense in rather big communities with a lot of different skills and experts. In such a context money may be a good medium to activate skills which would not be activated without it . Money also seems to be a way of regulating the demand for a good or a service, if no other form of regulation can be found. The crucial question is where money should be used and where it should not.

Communication environments demand for general accounting systems

"Essentially, communities may provide resources for the redress of infractions and forfeitures of debts that might not otherwise be redeemable. Social pressure from insult to incarceration to make good on all debts helps communities maintain the essential collective good of trust. The benefit of maintaining a generalised accounting system (one that allows for credit and does not demand intensive monitoring) is supported by experimental research (Kollock, 1992) in which it was found that generalised accounting systems yield much greater mutual benefit than tight systems that demand in kind exchanges at all turns." [4]

Small and closed communities seem to facilitate the maintenance of generalised accounting systems[1]. Closure and well-defined boundaries seem to be a good way to maximise mutual benefits and to avoid free-riding. Families benefit from the fact that family members are more or less defined by birth or adoption. The fraternal communities in monasteries and convents are also based on well defined rules of incorporation of new members; common goods only belong to members of the community and not to outsiders. If outsiders may benefit is up to the community. It seems to be paradox, but outsiders may profit from the strong boundaries that keep them out if the protected community is able to produce a surplus, because of its special form of co-operation and if that surplus is given to the outsiders. By blurring the boundaries, outsiders would also lose something in the long term.

In summary, there seems to exist a trade-off between boundaries and the need for money and other accounting systems: Strong boundaries allow mutual benefit in very generalised accounting systems. Weak boundaries increase the need for regulation and a rather strict accounting system.

In fact a good deal of experience gained in CMC shows that open discussions with a large number of participants tend to lose quality and social coherence. On the other hand, strong boundaries may not fit many conferences and discussions and may interfere with the open character of the Internet.

The model as a rating game

To give the better impression of how the co-operation in rather complex systems may be designed, the following *game* was invented. The term *game* is used because of the game-like rules. Rules can be seen as automatic relations between the amount of rating units someone has collected and his/her role in the system. Please note that it is not a game in the sense of game theory, but in the sense of party games. Nevertheless it alludes to real social settings and mechanisms. The game introduced here is one of thousands of possible games that may be designed for communities, and its main objective is to discuss the impact of different rules on the development of communities. To enable a better understanding of the game it was embedded in a real-life scenario:

Imagine a trans-disciplinary research context, for example `Artificial Intelligence'(AI) or `Human-Computer-Interaction'(HCI)[2]. In such contexts of knowledge production there is a rather rapid change of themes and opinion leaders. The role structure of a messaging-system should mirror that by allowing maximum flexibility in opening new discussions and in supporting a permanent flow of the members' degree of activity between being active and organising and being a rather passive consumer.

In other words, a person's roles and social functions may vary greatly. There are periods of high activity, when someone is eager to co-operate intensively, and others of lesser activity when someone prefers to lean back or to retire. The assignment of real-life roles therefore is more or less 'soft' and fluent.

The idea behind the following game is that teamwork very much depends on both a clear definition of roles and on the members' flexibility to change roles and to perform other functions. Hierarchy and the will to remain `on top' mostly interfere with common goals or impede defining such goals. Therefore a game with dynamic hierarchies is introduced. A number of rules prevent persons remaining on top and therefore becoming exhausted and obstructing others endowed with fresh forces. It may be assumed that people will accept rotation as something natural and positive for everybody.

Players may hold three different roles shown in the figure below.

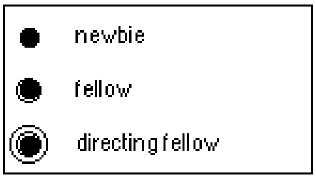


Figure 2: Players

Newbies are newcomers that become normal fellows after a certain time (according to rules to be defined later.) Fellows and directing fellows match the two modes of activities described above. Members are to rotate between the roles of a normal fellow and the role of a directing fellow. Figure 3 is to give an impression on how newbies, fellows and directing fellows may be grouped into several discussions.

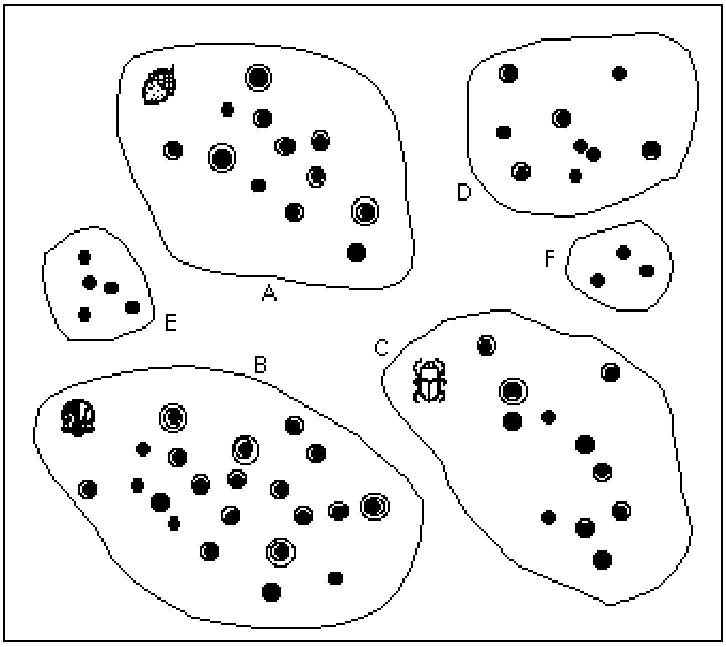


Figure 3: A Field Of Research With Different Discussions

Normally newbies start in open discussions dealing with very general and new topics (like discussions E and F in figure 3).Only newbies who have been personally invited by fellows or directing fellows may enter a closed discussion (like A, B or C in figure 3). Closed discussions deal with more specific topics. Fellows and directing fellows may enter all closed discussions.

The function of rating in the game context

In this context, rating performs two functions:

1. interesting and outstanding contributions are marked (orientation function)

2. people's roles change according to the amount of rating units they have collected by writing interesting and outstanding contributions (role-dynamic function).

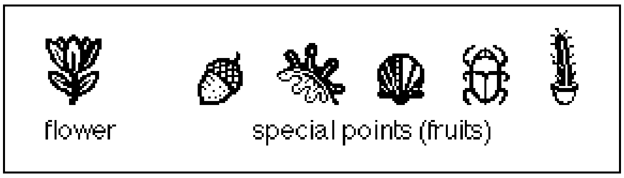


Figure 4: Rating Units

Figure 4 shows different ways of rating units to rate contributions. Everybody (also newbies) may rate with *flowers*. Special points named *fruits* are rating units reserved to directing fellows. *Fruits* express appreciation by a directing fellow and are a special honour to receive. E.g. directing fellows of discussion A in figure 3 may rate outstanding contributions with an acorn:

The different kinds of fruits belong to directing fellows in different discussions. The directing fellows of discussion B may only rate with shells. Directing fellows of discussion A use acorns. Acorns and shells, like any *fruit*, cannot only be awarded to contributions in discussion A, but to any contribution in the whole system, as there might be interesting ideas also outside a discussion. When discovering an acorn outside of discussion A, one knows that there is a very interesting idea with respect to the topics belonging to discussion A.

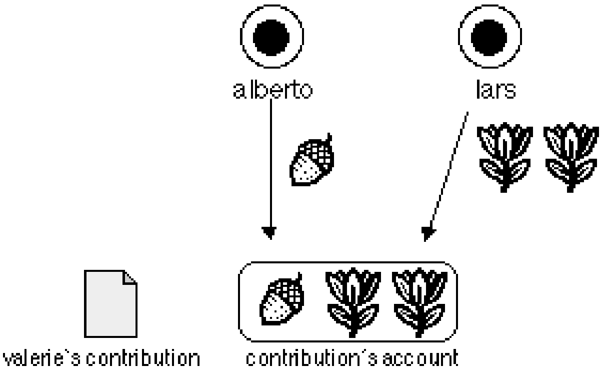


Figure 5: Accounting System (Part 1)

All ratings are collected in the contributions account. Figure 5 shows that different persons may of course rate the same contribution.

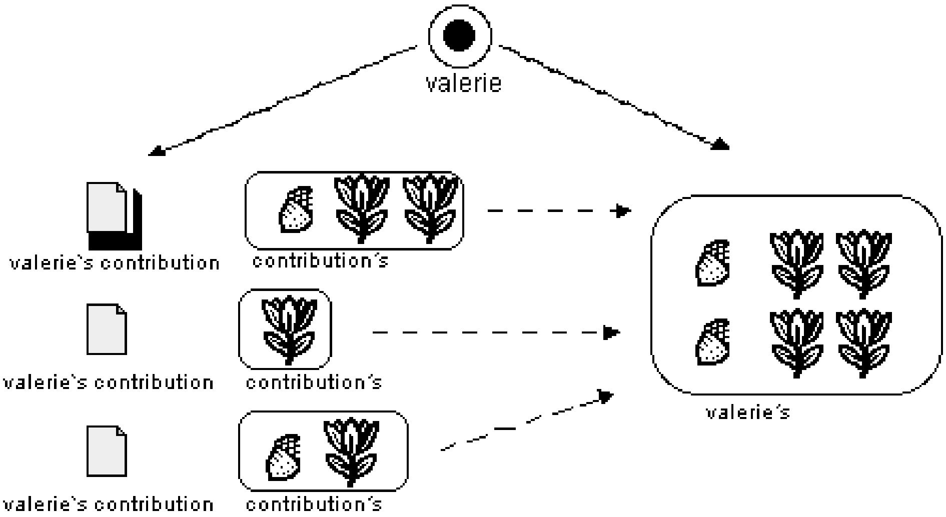


Figure 6: Accounting System (Part 2)

The sum of all ratings for all contributions of a certain person are collected in his or her personal account. The*contributions account* orients the reader about the quality of the content, the *personal account* is important for a person's status within the system.

Rule 1: One to three flowers can be awarded at once. Awarding flowers does not affect the donor's own account of flowers.

Role 2: One to three fruits may be awarded by directing fellows. Awarding fruits reduces the donor's account. Fruits of all kinds can be converted into each other. E.g. if a fellow from discussion A has obtained two shells, three acorns and one cactus, his account of fruits is six. Therefore he is allowed to give six acorns after becoming a directing fellow (because of rule 5) in discussion A.

How to become a fellow?

Rule 3: A newbie becomes a fellow by receiving any kind of fruit from a directing fellow.

Rule 4: A newbie becomes a fellow if he receives three flowers from at least two different persons.

How to become a directing fellow?

Rule 5: After obtaining three fruits of one kind, a fellow becomes a directing fellow in the respective discussion. Example: After obtaining three acorns a fellow becomes directing fellow in discussion A (in figure 3). It is possible to be a directing fellow in different discussions.

Rule 6: After spending all his fruits he returns to be an ordinary fellow.

These six basic rules are very simple, but nevertheless allow a great variety of possible behaviour and tactics by the players that need further discussion and additional rules, which are described in the following section.

Tuning the rating game

Getting and losing access

It is quite easy to become a fellow: The rating mechanisms defined in rules 3 and 4 construct a kind of boundary that keeps unwanted people out. This coincides with what was stated above, that boundaries may increase the mutual benefit of members. It is possible for a `gang' of three people to bypass rule 4 by awarding high ratings to each other. This may be avoided by special rules but such `gangs' may also be regarded to be inspiring and innovating. If their influence on the discussion is considered destructive, they may be thrown out by a special rule, e.g.

Rule 7a: Two directing fellows are enough to throw a fellow out of the respective discussion. If there is only one directing fellow, that one has the right to ban fellows. Their fellowship is not influenced by being banned.

This is a rather authoritarian rule allowing fast reactions to unwanted people. It could be changed into a more democratic rule, like

Rule 7b: The banning of people needs a voting procedure in which all members may take part. Simple majority is needed.

Directing discussions

It is quite easy to become a fellow but not so easy to become a directing fellow. As pointed out above, there should be special rules preventing a directing fellow from keeping that role for life. This may be acceptable in a traditional discipline but not in a context of very dynamic knowledge production. The rules already mentioned do not force directing fellows to spend their fruits. Therefore we have to define an additional one.

Rule 8 A directing fellow must spend at least three fruits within one month. Recipients have to be ordinary fellows.

This rule has two implications. First, it forces directing fellows to spend their fruits, and second it produces a new generation of directing fellows. Rule 8 is the key rule producing role rotation, as it is impossible to create a closed circuit of fruits among directing fellows only. One consequence of rule 8 might be that rotation only takes place among a limited number of persons (see figure 5).

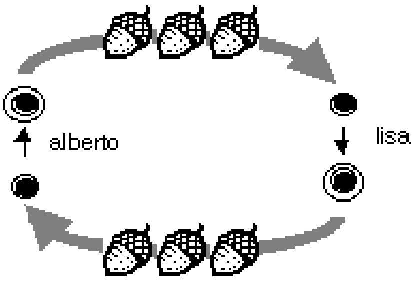


Figure 7: A Strategy Of Mutual Support

Alberto is rating Lisa's contributions with his fruits until she becomes a directing fellow. She rewards this by making him a directing fellow again. This kind of *rotation gang* may be avoided by additional rules, but this feature need not be counter-productive as it provides rotation anyway and is visible for everybody. People may complain about *rotation gangs*, or they may insist on a change in the rules, or simply leave the discussion. One effective measure against small rotation gangs could be an addition to rule 5:

Rule 5.1: To become a directing fellow one needs at least three fruits of one kind from at least two different directing fellows.

To bypass this rule one needs even more co-operation, which may be a wanted side-effect.

How new discussions emerge and evolve

New discussions may emerge for two major reasons:

1. Fellows and directing fellows try to keep newbies out. Therefore they set up an additional discussion which may compete with the original one. But this needs an additional rule like

Rule 10: It needs at least three persons to start a new discussion. Everyone, including newbies, may do so.

2. Another reason could be that there is a real need for new discussions, as new topics have arisen.

New discussions do not have directing fellows. They are more or less anarchic and they have to pass a phase of self-definition and self-construction until they become regular discussions. Examples are discussions D, E and F in figure 3. The change from provisional to regular discussions must be defined as rules.

Rule 11: After a turnover of at least 30 flowers within a month a discussion becomes regular. The two participants holding the largest number of flowers become directing fellows.

There might be another rule influencing the number of directing fellows. If there is a lot of activity in a discussion, a larger number of directing fellows may be needed:

Rule 12: If the number of contributions in a month exceeds 200, an additional directing fellow is nominated. It is the fellow with the largest number of flowers.

There are no restrictions for inventing new rules changing the system's behaviour. All participants may discuss intensively the impact of new rules. Procedures for the introduction of new rules may be introduced as well.

A rating game for Web4Groups

The limitations in the assignment of roles and `social functions' of Web4Groups require certain modifications. Like the above game, Web4Groups contains three roles:

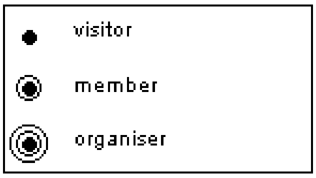


Figure 8: Roles in Web4Groups

The graphical representation of these roles allude to a functional resemblance between visitor and newbie, member and fellow, as well as organiser and directing fellow. These roles are more or less equivalent but certain differences exist: Membership in Web4Groups is limited to a certain discussion (Forum, Workgroup, etc.) whereas fellowship in the game is not. As a fellow you have the same rights and functions in all discussions. That is the reason why the game must be reduced to a single discussion. Nevertheless it is possible to combine rating and role assignment in a similar way as in the game above.

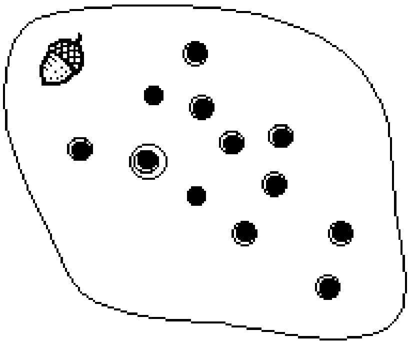


Figure 9: An Activity (Public Forum, Shared Workspace) in Web4Groups

Figure 9 shows that there is only one single organiser. Visitors are not registered, they may read and write but they have no personal account where rating units may be collected. Therefore, no rule can be invented which regulates how a visitor becomes a member. But visitors themselves may have the right to rate. In Web4Groups the main difference between a visitor and a member is registration. Nevertheless it is possible to provide for switching between the roles of member and of organiser. The respective rules are very similar to those in the game above. The main difference is that every member starts with an acorn on his account. Giving acorns is not reserved to organisers.

Rule 1: One to three flowers can be awarded at once. Awarding flowers does not affect the donors' own accounts of flowers. Flowers only serve for orientation and do not influence role assignment.

Role 2: One to three acorns may be awarded by the organiser. Every member starts with an acorn on his or her account. Awarding acorns affects the donor's account .

Rule 3: After obtaining four acorns, a member becomes the organiser of the discussion. The old organiser returns to being a normal member, but keeps his account of acorns. By default, the founder of a discussion is the organiser and has three acorns.

So it needs at least one person's support to become an organiser: Three acorns from the old organiser and a person's own acorn are enough. An additional rule is needed which will make people spend their rating units.

Rule 4: Organiser must spend an acorn every week. Normal members must spend an acorn every month.

Rule 5: Acorns that have not been spent are randomly distributed among people who have obeyed rule 4.

Rule 4 provides for the circulation of rating units. Rule 5 avoids an automatic decrease of rating units. If a person loses an acorn, another person must get one to keep the number of rating units constant.

Please note that there is no boundary keeping unwanted and completely uninformed people out. A group of four new members are able to make one of them an organiser.

Conclusion

We have shown that rating models shall include concepts of group dynamics and general value assignment and circulation to keep a virrtual group vital when exchanging messages. Existing rating transport tools such as PICS [5] can be used to visualize ratings, but they lack of means to gather ratings or assigning user roles. Rating collection interfaces and dynamic feedback mechanisms will be further investigated and tested with real users in the framework of the SELECT project.

Human beings tend to gather symbolic capital all their life long, with hope for immortality through a "place in history" or at least for being kept in mind within specific communities (Bordieu [7]). However, on the edge to a new millenium, chances to obtain symbolic immortality have reduced due to a highly differentiated society. This sociological insight from post-modern thinkers is applied with our model to the virtual society: no one's homepage can make somebody ever or well known, but useful activities for changing groups is a motivation and justification for being active in a virtual society.

References

[1] ALTON-SCHEIDL, Roland, Samba Diallo, Gernot Tscherteu, Social Functions in Virtual Communities; in: Informatik Forum, Vol. 10, Nr. 2 (http://www.Web4Groups.at/w4g/source/RateVoteSource.html), Vienna, 1996.

[2] GIBBONS, Michael et. al., The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies, London: Sage 1996.

[3] KOLLOCK, Peter & Marc A. Smith, 1994: Managing the Virtual Commons. Co-operation and Conflict in Computer Communities; Los Angeles: UCLA (http://www.sscnet.ucla.edu/soc/csoc/vcommons.htm), 1994.

[4] SMITH, Marc A., Voices from the Well: The Logic of the Virtual Commons; Los Angeles: UCLA (http://www.sscnet.ucla.edu/soc/csoc/virtcomm.htm), 1996.

[5] Platform for Internet Content Selection: http://www.w3.org/pub/WWW/PICS/

[6] BAUDRILLARD, Jean: Symbolic Exchange and Death.

[7] BORDIEU, Pierre: The Production of Belief: Contribution to an Economy of Symbolic Goods. Media, Culture and Society, 2, 1980: 261-293.

[8] PALME, Jacob: SELECT - Choices in the Implementation of Rating http://www.dsv.su.se/~jpalme/select/rating-choices.html

A Visual Tagging Technique for Annotating Large-Volume Multimedia Databases

A tool for adding semantic value to improve information filtering

(Post Workshop revised version, November 1997)

Konstantinos Chandrinos, John Immerkær, Martin Dörr, Panos Trahanias
Institute of Computer Science
FORTH
71 305 Heraklion, Crete
Greece
{kostel | jimm | martin | trahania}@ics.forth.gr

Abstract

The Computer Vision & Robotics Lab along with the Information Systems & Software Technology Lab at ICS-FORTH have recently developed a technique for marking-up different media to provide rapid entering of semantic information on large-volume multimedia databases. We expect this per medium semantic information to improve filtering of query results and facilitate retrieval of compact, context-sensitive information from large-volume multimedia databases. To that end, we present open design issues. Our technique offers the ability to interactively mark-up parts of the representation on the medium of choice (e.g. areas on a document image) and associate local information with text, images, sound-clips, video or subparts of them without altering the original. The output is XML files that can be parsed by viewers to navigate through the semantic information. The XML files can also be used by intelligent search agents for different media in order to refine the accuracy of retrieval in a query. The basic principles of implementation allow for database-flexibility through the use of XML. Developing all the software in Java allows for net-centric administration and use of a number of heterogeneous databases through standard Internet browsers. The Java implementation also guarantees platform-independence. This work has been developed under the ARHON project and is part of an ongoing effort to answer the need for novel ways of data entering and registering in large multimedia databases. The methods presented are intended to carry over the analogy of text mark-up by a database administrator or approved user, to other media such as still images, video and sound.

Introduction

In recent years ICS-FORTH has designed, implemented and installed a number of large-volume multimedia databases. On launching the ARHON project (A Multimedia System for Archival, Annotation and Retrieval of Historical Documents, Jan. 1997 - June 1998) we were faced with a good 1,500,000 archive manuscripts from the late 1600s to early 1900s in varying degrees of deterioration. The ARHON project [1] is to provide the framework under which these documents can be turned into a digital library accessed through an intranet at first, and then over Internet. As a test bed for a model we work on a sub-archive of about 100,000 documents. These documents have to be scanned to an electronic form, processed for image correction and then transcribed or even translated as some of them are in a foreign language compared to that of potential users.

The problem analysis

A first approach would be to combine OCR technology with text-based mark-up (e.g. SGML). However, the inaccuracy of state-of-the-art OCR software on all but the most modern and uniformly printed documents (let alone manuscripts) still demands thorough proof-reading and error correction. Also, the intensive manual labor and financial cost required for SGML text mark-up were deemed as barriers to our purpose. Even if one could get a perfect OCR result on 17th century manuscripts so as to apply Full Text Indexing techniques, the most important view of the resulting digital library, the semantic information would be missing. As a result additional mark-up would be necessary and in our case would have to be carried out by specialists who first would have to become familiar with SGML specific tools. As a last argument against textual mark-up we considered that it is a very invasive procedure with respect to the raw data, since textual mark-up changes the original. Even worse, in our case we have to cater for multiple annotations for the same document coming from several researchers. The above situation gives rise to a number of questions:

- *Accommodating different levels of authority*. Scholar researchers come at different levels of authority as well as different fields. Consulting the login data, assigning different weights to scholars as a default, according to the standing estimations of the archivists, seems an acceptable choice. However, authorities seem to “decay” in time as new scholars with new interpretations show up. This indicates that a proper weighting system had better be a function of the utilisation of annotations by other users. The users of the digital archive should be allowed to change such system rating of scholars either per session or through personalised profiles.

- *Ranking the query results according to relevance*. This could be done using the number of annotations for each result. Priority should be given to images or other more complex media (e.g.

video, sound) since these present higher fidelity to the originals as opposed to transcribed ASCII text.

Our approach

To overcome the shortcomings of OCR and SGML manual markup we have designed, developed and implemented a novel approach for visual tagging. In our approach we encourage the user to mark-up parts of the representation on his/her medium of choice (e.g. areas on a document image) and associate local information (e.g. filename, spatial or temporal co-ordinates etc.) with text, images, sound-clips, video or subparts of them without altering the original. To illustrate our approach we present here ***ImageTagger***, an implementation of our ideas that focuses on image documents (cf. Fig 1) and associates selectable faceted keywords to them. Nevertheless, the philosophy of our design and implementation is valid for any other media.

This mark-up technique produces a file in XML (eXtensible Markup Language, [2]) format that can be parsed to update a varying range of databases through the appropriate Java Database Connectivity driver (JDBC, [3]). Currently, we are implementing a JDBC for the Semantic Index System (SIS, [4]), a system developed at FORTH for describing and documenting large evolving varieties of highly interrelated data, concepts and complex relations. Since the format of the produced files is open, implementations for any other database system, whether RDBMS or object-oriented is considered trivial. In Figure 1 we also demonstrate a labelling technique implemented in our viewer. The viewer parses the output file to create individual labels for marked regions.

Using JavaServer 1.0 as an http server and servlet technology [5], we are now testing simultaneous media based queries. With the innate multithreading capability of servlets and Java this can be done without significant programming overhead.

System evaluation

What we consider of significant importance is that this design allows databases of different technology, such as a thesaurus and an RDBMS, to be updated through parallel parsing of the XML file. Additionally, implementation in Java allows the same code to be executed on different platforms. Currently Java ver. 1.1 is supported by Windows 95/NT and Solaris. To ease installation and maintenance and cater for repositories that are not implemented in these two platforms we have built into our program the ability to execute as a Java applet inside a standard Internet browser (cf. Fig. 1 presenting a snapshot of the ***ImageTagger*** running inside Netscape Communicator 4.03). With this last kind of use there is no client software installation and the user always gets the latest version.

Future directions

Next in our plans comes the implementation of taggers for other media, namely sound and video and the integration of all such media taggers using existing Web technology (HTML and Java) combined with the latest evolution in mark-up languages (Synchronised Multimedia Integration Language, SMIL [6]). In order to test query refinement with these tagging techniques we intend to apply search agents for each medium that will execute the query on their part and filter the results. Points we shall investigate in this direction include:

- *Automation of the search and filtering process through the use of intelligent agents*. Autonomous media-specific agents can be used to execute the actual search and collaboratively filter the results. Hence the agents will be required to reason about the use of the resources and negotiate among them in an agreed protocol. Provision should be taken for the machinery of agent-creation to allow further production of agents according to the expansion of the digital library. There is definitely a cost in the negotiation part of the agents function which even in the best design architecture is inversely proportional to the available bandwidth.

- *Creation of a user profile*. This profile will be published along with the query to the appropriate agents so as to guide them through the search process. We share the view that casual users should be given less retrieved information and a “More like this” option, compared to expert users that can pinpoint what they are looking for and can identify it easier in fine detail. Additionally, similarities in past queries from the same or a different user may provide stimulation for query refinement. Searching through the profile base poses an extra overhead in implementation and real time efficiency. In this direction we intend to implement some directory service solution, based perhaps on LDAP [7].

References

[1] K.V. Chandrinos, J. Immerkær and P.E. Trahanias, ARHON: A Multimedia Database Design for Image Documents, submitted to EUSIPCO 98, Special Session on Multimedia Signal Processing

[2] http://www.w3.org/xml/

[3] http://www.javasoft.com/products/jdbc/

[4] http://www.ics.forth.gr/proj/isst/Systems/SIS

[5] http://www.javasoft.com/products/java-server/

[6] http://www.w3.org/smil/

[7] Joao Ferreira, Jose Luis Borbinha, Jose Delgado: Using LDAP in a Filtering Service for a Digital Library, 5th DELOS Workshop, Budapest, Nov. 1997

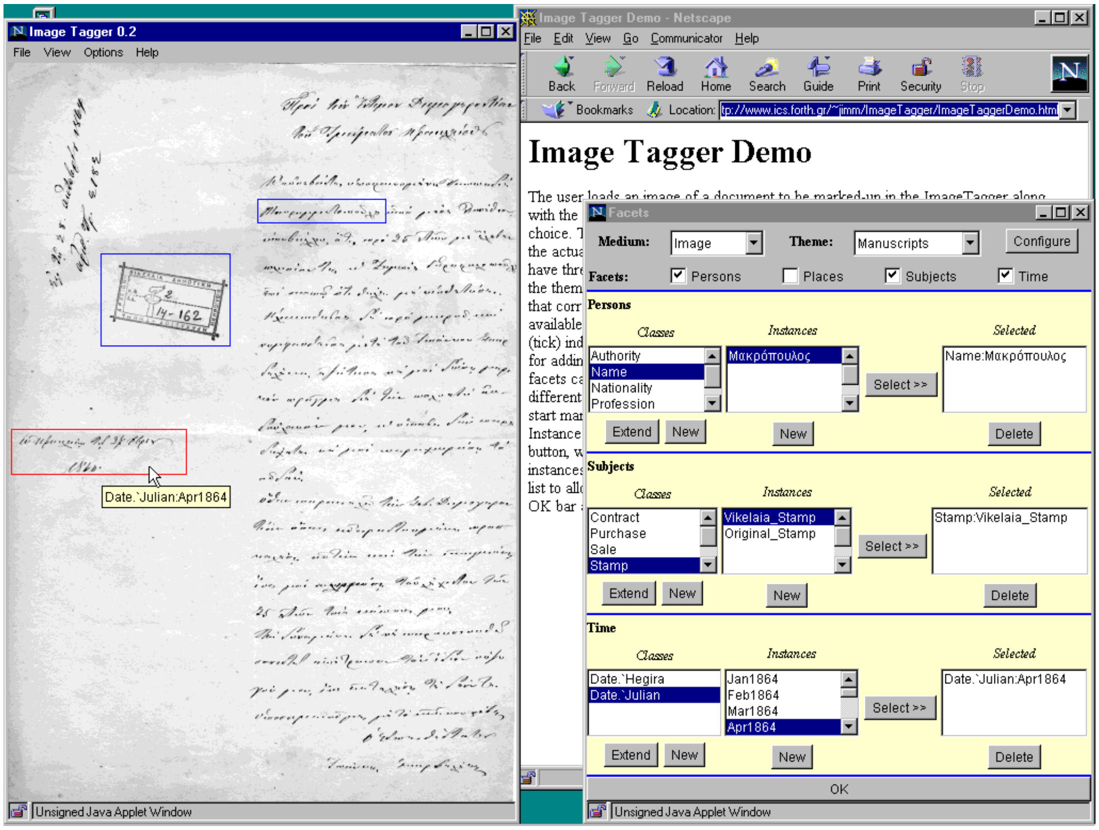


Image Tagger 0.2
File View Options Help
Image Tagger Demo - Netscape
File Edit View Go Communicator Help
Back Forward Reload Home Search Guide Print Security Stop
Bookmarks
Facets
Configure
Persons
Classes
Instances
Selected
Extend New New Select >> Delete
Subjects
Time
OK
Unsigned Java Applet Window

Date.`Julian:Apr1864

Location: tp://www.ics.forth.gr/~jimm/ImageTagger/ImageTaggerDemo.html

Image Tagger Demo

The user loads an image of a document to be marked-up in the ImageTagger along with the … choice. T… the actua… have thre… the them… that corr… available… (tick) ind… for addin… facets ca… different… start mar… Instance… button, w… instances… list to allo… OK bar …

Medium: Image   Theme: Manuscripts

Facets: ☑ Persons □ Places ☑ Subjects ☑ Time

Persons: Classes: Authority, Name, Nationality, Profession — Instances: Μακρόπουλος — Selected: Name:Μακρόπουλος

Subjects: Classes: Contract, Purchase, Sale, Stamp — Instances: Vikelaia_Stamp, Original_Stamp — Selected: Stamp:Vikelaia_Stamp

Time: Classes: Date.`Hegira, Date.`Julian — Instances: Jan1864, Feb1864, Mar1864, Apr1864 — Selected: Date.`Julian:Apr1864

Figure 1. ImageTagger demo running as an applet in Netscape Communicator

The DELOS Working Group is a part of the ESPRIT Long Term Research Programme (LTR No. 21057) and is managed by ERCIM.

The DELOS Partners are:

ERCIM members:

CLRC, CWI, CNR, FORTH, GMD, INRIA, INESC, SICS, ETH-SGFI, SINTEF Telecom and Informatics, MTA SZTAKI, VTT

Non-ERCIM members:

University of Michigan, USA
Elsevier Sciences, The Netherlands

For additional information, please contact

Costantino Thanos
Istituto di Elaborazione della Informazione, Consiglio Nazionale delle Ricerche
Via Santa Maria 46
I-56126 Pisa
Tel: +39 50 593492, Fax: +39 50 554342, E-mail: thanos@iei.pi.cnr.i

DELOS web site: **http://www.area.pi.cnr.it/ErcimDL/delos.html**

ISBN 2-912335-04-3

The European Research Consortium for Informatics and Mathematics (ERCIM) is an organisation dedicated to the advancement of European research and development, in the areas of information technology and applied mathematics. Through the definition of common scientific goals and strategies, its national member institutions aim to foster collaborative work within the European research community and to increase co-operation with European industry. To further these objectives, ERCIM organises joint technical Workshops and Advanced Courses, sponsors a Fellowship Programme for talented young researchers, undertakes joint strategic projects, and publishes workshop, research and strategic reports as well as a newsletter.

ERCIM presently consists of fourteen research organisations from as many countries:

Central Laboratory of the Research Councils

Rutherford Appleton Laboratory
Chilton, Didcot
GB-Oxon OX11 0QX

Tel: +44 123582 1900
Fax: +44 1235 44 5385
http://www.cclrc.ac.uk/

Centrum voor Wiskunde en Informatica

Kruislaan 413
NL-1098 SJ
Amsterdam

Tel: +31 205929333
Fax: +31 20 592 4199
http://www.cwi.nl/

Consiglio Nazionale delle Ricerche

IEI-CNR
Via S. Maria, 46
I-56126 Pisa

Tel: +39 50 593 433
Fax: +39 50 554 342
http://bibarea.area.pi.cnr.it/ERCIM/welcome.html

Czech Research Consortium for Informatics and Mathematics

FI MU
Botanicka 68a
602 00 Brno

Tel: +420 2 6884669
Fax: +420 2 6884903
http://www.utia.cas.cz/CRCIM/home.html

Danish Consortium for Information Technology

CIT
Forskerparken
Gustav Wieds Vej 10
8000 Århus C

Tel: +45 8942 2440
Fax: +45 8942 2443
http://www.cit.dk/ERCIM/

Foundation of Research and Technology – Hellas

Institute of Computer Science
P.O. Box 1385
GR-71110 Heraklion, Crete

Tel: +30 81 39 16 00
Fax: +30 81 39 16 01
http://www.ics.forth.gr/

GMD – Forschungszentrum Informationstechnik GmbH

Schloß Birlinghoven
D-53754 Sankt Augustin

Tel: +49 2241 14 0
Fax: +49 2241 14 2889
http://www.gmd.de/

Institut National de Recherche en Informatique et en Automatique

B.P. 105
F-78153 Le Chesnay

Tel: +33 1 39 63 5511
Fax: +33 1 39 63 5330
http://www.inria.fr/

Instituto de Engenharia de Sistemas e Computadores

Rua Alves Redol 9
Apartado 13069
P-1000 Lisboa

Tel: +351 1 310 00 00
Fax: +351 1 52 58 43
http://www.inesc.pt/

Swedish Institute of Computer Science

Box 1263
S-164 28 Kista

Tel: +46 8 752 1500
Fax: +46 8 751 7230
http://www.sics.se/

Schweizerische Gesellschaft zur Förderung der Informatik und ihrerAnwendungen

Dept. Informatik
ETH-Zentrum
CH-8092 Zürich

Tel: +41 1 632 72 41
Fax: +41 1 632 11 72
http://www-dbs.inf.ethz.ch/sgfi/

Stiftelsen for Industriell og Teknisk Forskning ved Norges Tekniske Høgskole

SINTEF Telecom & Informatics
N-7034 Trondheim

Tel :+47 73 59 30 00
Fax :+47 73 59 43 02
http://www.informatics.sintef.no/

Magyar Tudományos Akadémia – Számítástechnikai és Automatizálási Kutató Intézete

P.O. Box 63
H-1518 Budapest

Tel: + 361 166 5644
Fax: + 361 166 7503
http://www.sztaki.hu/

Technical Research Centre of Finland

VTT Information Technology
P.O. Box 1200
FIN-02044 VTT

Tel:+358 9 456 6041
Fax :+358 9 456 6027
http://www.vtt.fi/

ERCIM
Domaine de Voluceau, Rocquencourt, B.P. 105, F-78153 Le Chesnay Cedex, FRANCE
Tel: +33 1 39 63 53 03 Fax: +33 1 39 63 58 88