TOWARD SOCIAL SEARCH
From Explicit to Implicit Collaboration to Predict Users' Interests

Luca Longo, Stephen Barrett and Pierpaolo Dondio
Department of Computer Science and Statistics, Trinity College Dublin

Keywords: Social search, User behavior, Computational trust, Web site classification.

Abstract: The concept of *social search* has been acquiring importance in the WWW as large-scale collaborative computing environments have become feasible.This field focuses on the reader's perspective in order to assign relevance and trustworthiness to web pages. Although current web searching technologies tend to rely on explicit human recommendations, these techniques are hard to scale as feedback is hard to obtain. Implicit feedback techniques, on the other hand, can collect data indirectly. The challenge is in producing implicit web-rankings by reasoning over users' activity during a web-search without recourse to explicit human interventions. This paper presents a comparison between explicit and implicit users' feedbacks upon web pages. An experiment, involving 25 volunteers explicitly evaluating the usefulness of 12 thematic web-sites, was performed implicitly gathering their web browsing activity. The results obtained prove the existence of a strong correlation between explicit judgments and generated implicit feedbacks.

1 INTRODUCTION

The main advantage of systems supporting *social search* is that Web pages are considered relevant and trustworthy from the reader's perspective rather than web sites owners. Such solutions contrast with the majority of the current search engines, above all Google, whose Page-Rank algorithm assigns importance to web-pages based on the analysis of the link structure of the Web. A key open challenge in designing *social search* systems is to improve the overall information seeking and consuming activities on the Web. Reading time, scrolling, cut-paste are all considered relevant implicit sources of user preferences (Kelly and Belkin, 2001). In the current web-searching technologies, a substantial inhibitor in gathering this content from users is that they tend to be resistant to invasive techniques and there exists lack of motivation to actively generate recommendations. Implicit feedback techniques gather data indirectly and the key issue is to produce implicit web-rankings automatically deduced by reasoning on the activity performed by users over web-pages. In this work we propose a novel approach to collaborative *social search* that analyses users' actions during Internet sessions capturing this activity. Such activity embodies implicit 'human judgement' where each web-page has been viewed and endorsed by one or more people concluding that it is There are three key benefits of such a solution. First, as each result has been selected by users, by reasoning on their behaviour, it is possible to obtain a relevant degree of trustworthiness. Second, the *social search* engine operates concurrently over continuously updating of user activity and so it is well positioned to display stronger results more current or in context with changing information. Third, it is possible to reduce the impact of link Spam by relying less on link structure of web-pages. The paper is structured presenting in 2 related works; in 3 we underline the hypothesis, we describe the experiment and the formal model commenting obtained results. We conclude in 4, with future work and open issues.

2 RELATED WORK

The concept of *social search* has been acquiring importance as the World Wide Web grows in size and web-searching technology has become an essential need in web-browsing. Several methods have been conceived from the simplest, based on sharing bookmarks, to more sophisticated approaches that combine human intelligence with computer paradigms (Agichtein, et al,.2006) supporting collaborative gath-

Longo L., Stephen B. and Dondio P.
TOWARD SOCIAL SEARCH - From Explicit to Implicit Collaboration to Predict UsersâĂŹ Interests.
DOI: 10.5220/0001841406840687
In *Proceedings of the Fifth International Conference on Web Information Systems and Technologies (WEBIST 2009)*, page
ISBN: 978-989-8111-81-4

ering and collaborative directories. In (Atterer et al.,2006) the authors focused on tasks such as classifying the user with regard to computer usage proficiency or making a detailed assessment of how long it took users to fill in fields of a form. They developed an HTTP proxy that collects data about mouse movements, keyboard input and more. Similarly in the work of Velayathan and Al. (Velayathan and Yamada, 2007), an unobtrusively framework logs and analyses users' behaviour to extract effective rules to evaluate web-pages using a machine-learning techniques. In the work reported in (Kelly and Belkin, 2001) the authors focused on the hypothesis that users will spend more time, scroll more often and interact more with those documents they find relevant. Similarly, in the work of Weinreich et Al., (Weinreich et al., 2006) authors found that users spend less than 12 seconds on nearly 50% of the web-pages shown to them demonstrating users make nearly 50% of their decision to navigate to the next page before reading substantial part of the contents.

Collaboration is a process where people interact each other toward a common goal, by sharing their knowledge, learning and building consensus. There are two main way to provide judgement: explicitly and implicitly. In the former way, users can provide feedback using a specific metric, for example as in eBay and Amazon community. In the latter way, implicit judgements are inferred from user behaviour while doing a specific action. Collaboration applied to the Web 2.0 supports a new kind of shared intelligence, named *Collective Intelligence* where users are able to generate their own content building up an infrastructure where contributions are not merely quantitative but also qualitative.

The relevance of Trust and Reputation in human societies is indisputably recognised A trust-based decision is a multi-stage process on a specific domain. This process starts identifying and selecting pieces of trust evidence, generally domain-specific, conducting an analysis over the application involved. Subsequently, trust values are produced performing a Trust computation over the pieces of evidences estimating the trustworthiness of entities in the domain considered. Both the previous steps are informed by a notion of trust in the Trust model and the final Trust decision is taken by considering the computed valued along with exogenous factor like disposition or risk assessments. The proliferation of collaborative environments represent good examples in which *Computational Trust* paradigms are applied in order to evaluate the trustworthiness of virtual identities. Longo et al. (Longo et al., 2007) conceived a set of rare trust evidences based on time and applied on Wikipedia, demonstrating how plausible Trust decisions can be reached using exclusively temporal factors. Teamwork and co-operation (Montaner et al., 2002) represent other areas where the game theory is the predominant paradigm considered to design *Computational Trust* models.

3 IMPLICIT/EXPLICIT COLLABORATION: EXPERIMENT AND TRUST MODEL

The *hypothesis* behind this work is to understand whether, taking into account an entity and applying *Computational Trust* paradigms by using reasoning techniques, explicit human judgements are correlated with the corresponding implicit derived feedback. We explore this question in the context of web-page media. If the answer is positive, i.e., there exists a correlation between them, it is possible to build up a collaborative environment achieving good predictions in a non-invasive way. In particular, we can conclude that, examining users' behaviour while surfing the Internet, we can generate a set of ranked results where the top ones represent the most valuable content considered by users and thus, by implication valuable to other similar users. We refer at this kind of collaboration as *implicit collaboration* to distinguish from the classic, *explicit collaboration*, where users expressly provide feedback, evaluations and judgements. Our solution was to log all the activity in the browser gathering the main events(E_i) that may occur during an Internet session. The logger does not perform any kind of computation, it does not apply any *Computational Trust* paradigms nor does it filter out events.

We conducted experiments in order to investigate the ability of our approach to gather logs of user behaviour. 25 unpaid volunteers, with different backgrounds, were recruited to participate in this study. We asked each of them to organise a trip to Morocco, 2-weeks long, surfing a pre-defined list of web-sites from which it is possible to collect information about popular cities, transports, hotels. We proposed a list of 12 selected urls, that users can use within 60 minutes in which they have to naturally interact with the browser, collecting useful data, cutting and pasting relevant information, bookmarking interesting pages, submitting data, saving picture or documents in order to recover this information in the future. Finally, we ask to each of them to explicitly provide a judgement of the usefulness of each web-site using a common scale from 1 to 10 (1 means not useful and 10 means

Table 1: Event hierarchy with associated weights.

♯	Type	Description	Weight
1	*very rare*	E_1: save as (page)	W_1= 22%
2	*rare*	E_2: bookmark	W_2= 18%
3	*rare*	E_3: printing	W_3= 18%
4	*not frequent*	E_4: save as (picture)	W_4= 10%
5	*not frequent*	E_5: download	W_5= 10%
6	*frequent*	E_6: cut & paste text	W_6= 8%
7	*frequent*	E_7: search text	W_7= 8%
8	*very frequent*	E_8: form input	W_8= 4%
9	*very frequent*	E_9: scrolling	W_9= 2%

very useful). In this experiment we assume that volunteers act not maliciously, hence the data contained in the logs file is the proper representation of the real actions performed by them while surfing the Internet. We assume also that volunteers do not change their behaviour in order to alter generated logs data. At the end of the experiment, a set of noisy information is obtained for each user, containing his activity while surfing the given web-sites. Since volunteers can jump from one given web-site to another one, gathered logs need to be filtered and aggregated to produce a well defined set of data that we refer as 'user-behaviour pattern'. For these reasons, we developed a function named 'filter/aggregator' that analyses the data, filters all the urls not in the given set and aggregate it grouping per web-page. This module produces 12 'user-behaviour patterns' containing the occurrences of events, one for each given websites. At this stage, it is now possible, by applying *Computational Trust* paradigms, to generate a unique value of a given pattern indicating the usefulness of a given web-site for a specific user. More than one *Computational Trust* factors may be adopted and each of their outputs may be aggregated to obtain a more precise usefulness/trustworthiness degree for a given web-site. Considering the 'filter/aggregator' function, 12 unique values, one for each given web-site are generated. We defined a basic *Computational Trust* model that extracts the occurrences of events in the 'user-behaviour pattern' and compute a real value in the range $[0..1]$. Since 'scrolling' events are more frequent than 'save as' events or a 'cut & paste' should be less important than 'bookmark' events, a hierarchy is needed to discriminate their importance. The goal of this paper is not to study a hierarchy of such events hence we consider the work presented by (Velayathan and Yamada, 2007), in which the authors provide the frequency of the most and the least events performed in their experiment while surfing the Internet, to propose our 'event-hierarchy' with weights, as described in the table 1.

The final real value of the model is computed by aggregating the occurrences of each event contained in the 'user-behaviour pattern' and by using the 'event-hierarchy' as shown in the following formal model:

$$Trust_{value} : BP[] \rightarrow [0..1]$$

$$Trust_{value}(BP[]) = \sum_{i=1}^{n} S(BP[E_i])$$

$$S = \begin{cases} 0 & \text{if } \forall E_i,\ BP[E_i] = 0 \\ \frac{1}{4}W_i & \text{if } BP[E_i] = 1 \ \&\ E_i = 6,7,8,9 \\ \frac{1}{2}W_i & \text{if } BP[E_i] = 2 \ \&\ E_i = 6,7,8,9 \\ W_i & \text{if } BP[E_i] \geq 2 \ \&\ E_i = 6,7,8,9 \\ \frac{1}{2}W_i & \text{if } BP[E_i] = 1 \ \&\ E_i = 4,5 \\ W_i & \text{if } BP[E_i] \geq 2 \ \&\ E_i = 4,5 \\ W_i & \text{if } BP[E_i] \geq 1 \ \&\ E_i = 1,2,3 \end{cases}$$

where BP is the 'user-behaviour pattern' vector, E_i is a specific event and n is the cardinality of the possible events. For *frequent* and *very frequent events*, the model assigns full corresponding weight (W_i), if and only if events occurred more than twice, otherwise $\frac{1}{2}$ of weight is returned for 2 occurrences, $\frac{1}{4}$ for just one and 0 otherwise. For *not frequent* event two occurrences are enough to set full weight and $\frac{1}{2}$ for just one occurrence. Eventually, for *rare events* such as 'bookmarks', 'save as page' and 'printing', just one occurrence is relevant, hence full corresponding weight is assigned. Taking into account both the 12 values produced by using our schema and the 12 judgements provided by each volunteer, we can use statistical correlation indexes to test the hypothesis. In particular, in this work we adopt the *Pearson's correlation coefficient* that measures the strength of the linear dependence between the implicit and the explicit values. If the correlation value obtained by considering the implicit value and the explicit judgement, for a given web-site and a given volunteer, tends to 1, a linear equation describes the relationship positively with the implicit value increasing with the explicit value. A score of -1 shows the inverted relationship and a value tending to zero shows that there is no linear relationship so the variable considered are independent. In order to test the hypothesis, we expect high correlation values, one for each use: if the majority of these values tend to 1, our hyphotesis is confirmed and we can sustain there exists a strong relationship between implicit feedback and explicit human judgements as captured by our model.

The set of Pearson's correlation values obtained from the experiment are encouraging. The mean of the users is 0.6 and more than 50% of them are above the value of 0.7. The 24% of them has a correlation value less then the threshold of 0.5. The 12%

of user show a low relationship between the explicit judgement and the implicit derived value: this fact shows our method did not succeed for 3 people. In 2 cases the experiment returns negative correlation coefficients, hence implicit and explicit values have an inverted relationship. 68% of users exhibit a correlation value above the mean. The strategy proposed is not strong in the short term: a web-site visited by a couple of users may have an average of trustworthiness higher than a web-site visited by thousand of users. An approach to resolve this problem may be the adoption of a threshold, explicitly defined or learned with unsupervised techniques, indicating the minimum number of users who had to have implicitly viewed the same site.

4 CONCLUSIONS AND FUTURE WORK

In this study we performed a context-dependent comparison between explicit human judgements, provided by volunteers, and implicit judgements derived by using *Computational Trust* techniques. Through an experiment we demonstrated how, taking into account a digital entity as a web-site, human explicit judgement can be strongly connected to the implicit derived value on the same entity. The evaluation was conducted by considering 12 Urls evaluated by users explicitly providing a degree of usefulness. During browsing sessions we logged users' activity and a behaviour-pattern, containing the occurrences of generated events, was extracted for each of them and each web-site. *Computational Trust* paradigms helped us to automatically evaluate these patterns and to generate trustworthiness values. The Pearson's coefficient was used to study the correlation between explicit users' judgements and derived usefulness values. Even considering a small number of users and a basic Trust model, encouraging results were obtained proving our hypothesis and underlying how it is possible to automatically evaluate entities such web-pages, by reasoning on users' activity over them. This is a new approach and thus there is further work to do. Authors believe it represents a start point to predict users' interests and to build up a third-generation *social search* engine based on *implicit collaboration*. Future works will be focused on experiment in malicious environments taking into account privacy/anonymity issues. New reasoning techniques should be investigated to better evaluate web-pages and new algorithms are needed to semantically connect searching queries to relevant sets of Urls generated by our schema.

REFERENCES

Kelly D. and J. Belkin N. (2001). Reading Time, Scrolling and Interaction: exploring Implicit Sources of User Preferences for Relevance Feedback During Interactive Information Retrieval, SIGIR 2001, New Orleans, LA, USA.

Agichtein E., Brill E. and Dumais S. (2006). Improving Web Search Ranking by Incorporating User Behavior Information. SIGIR 2006, USA.

Atterer R., Wnuk M. and Schmidt A. (2006). Knowing the User's Every Move - User Activity Tracking for Website Usability Evaluation and Implicit Interaction. WWW 2006, Edinburgh.

Velayathan G. and Yamada S. (2007). Behavior-based Web Page Evaluation, WWW 2007, Canada.

Longo L., Dondio P. and Barrett S. (2007). *Temporal Factors to evaluate trustworthiness of virtual identities*. IEEE SECOVAL 2007, France.

Weinreich H, Obendort H., Herder E. and Mayer M. (2006). Off the Beaten Tracks: Exploring Three Aspects of Web Navigation. WWW 2006.

Montaner M., Lopez B. and De La Rosa J. (2002). *Developing Trust in Recommender Agents*. AAMAS-02, Bologna, Italy, 2002.