A Reputation-based DSS: the INTEREST Approach

Donato Barbagallo[a],
Saverio Bruno [a],
Cinzia Cappiello [a],
Chiara Francalanci,
Maristella Matera [a], and
Lorenzo Radice [a]

[a] Politecnico di Milano - Dipartimento di Elettronica e Informazione
P.zza Leonardo da Vince, 32 -20134 - Milano - Italy
{barbagallo, cappiello, francala, matera}@elet.polimi.it

Abstract

Web 2.0 technologies have enabled an active role of the users, who can create and make available their contents very easily. This allows people to express their opinions, and distribute them through several means (e.g., forums, blog posts, social networks, etc.), thus increasing the amount of information on the Web. This high availability supports the users in their searches, but also raises the risk to find incorrect, ambiguous, or contradictory data. This paper proposes a platform for the construction of self-service environments for the selection and composition of trustworthy services for information access.

Keywords: Reputation; Quality Assessment; Mashups.

1 Introduction

Web 2.0 technologies enable an active role of the users, who can create and make available their contents very easily (Murugesan, 2007). This allows people to express their opinions, and to distribute them through several means, such as forums, blog posts and comments, social networks, etc. Of course, the availability of such an amount of information raises a set of issues concerning the research, the selection, and the representation of trustworthy sources and services. The information retrieved on the Web is indeed often characterized by inconsistent, incomplete, and not correct data, and users are sometimes not able to distinguish the right or the most suitable data along their needs. Furthermore, all the accessible sources could be better exploited if they wer combined so that to obtain a tangible added value. To respond to the previous needs, this paper discusses a novel approach, INTEREST (INnovaTivE solutions for REputation based self-Service environments), for the construction of personalized self-service environments through which users can build their view over the Web information space, by integrating trustworthy services for information access.

To understand the idea, let us consider a usage scenario in the health, wellness, and cultural tourism domain, a tourism segment which is gaining momentum. We suppose that an individual is in a trip and has a few hours to spend in a given city. He/she will

be searching through a mobile device for advice on the main attractions in town, based on his/her time frame and position. The platform should be able to provide suggestions. This raises a number of issues. First, the individual may wish to complement standard touristic information with Web 2.0 information from online communities. Since there exist alternative communities, he/she needs to choose the most dependable, active, and, possibly, authoritative community prior to start interacting with the other participants. Eventually, he/she would like to have a way to integrate different answers from multiple communities and bookmark the most reliable sources of information on specific subjects. For example, he/she has identified key people within different forums and would like to be informed whenever they contribute to the forum. Besides accessing "social" information, relevant information can also come from a network of sensors in charge of monitoring a variety of environmental properties (weather conditions, quality of water, etc). Users need to access such data through ad-hoc web services.

To our knowledge, there is a lack of works able to cover the above requirements exhaustively. Some work has been devoted to the trust of Web resources (Gil and Artz, 2009), focusing on content and making a distinction between *content trust* and *entity trust*. The literature is also oriented towards the interaction of Web Services (Malik and Bouguettaya, 2009), i.e., to the assessment of quality and service retrieval (Medjahed and Bouguettaya, 2005). Trustworthiness on the Web is often identified with popularity: this equation led to the success of the PageRank algorithm (Brin and Page, 1999) even if it does not necessarily conveys dependable information since highly ranked Web pages could be spammed. To overcome this issue, new algorithms are based on *hub and authority* mechanisms in the field of Social Network Analysis (SNA) (Kleinberg, 1999). Especially when considering services such as forums, in our approach we assume that it is important to evaluate even a single contribution: SNA can be used to evaluate each author's trustworthiness (Skopik et al., 2009).

The remainder of this paper is organized as follows. Section 2 illustrates a scenario that clarifies the novel requirements addressed by our research. Section 3 describes the architecture of the platform currently under development, while Section 4 discusses preliminary empirical results. Finally, Section 5 outlines our future work.

2 Platform Architecture

A scenario like the one illustrated in the previous section highlights two fundamental requirements: (i) to select dependable services that can fulfil specific information needs and quality requirements and (ii) to provide users with tools to compose on-demand personalized applications by means of the selected dependable services. Having these goals in mind, we designed INTEREST, a platform providing self-service mashup functionalities (Barbagallo et al., 2009), to help people create a personalized Web access environment. Based on a user profile, the platform searches for the "best" information sources, selects relevant and authoritative information, and then wraps it as data services supplying information about the reputation of given

entities. Such reputation services are then mashed up with other services of interest to create a personalized environment.

Figure 1 illustrates the main components of the proposed platform. The *Service Registry* stores a catalogue of services that can be used for the creation of the personalized user environment. They can be *data services*, providing a binding with a dependable data source through a description of the data structure, generic Web services enabling the retrieval of some relevant information, or also mashup APIs. The registry is populated by a domain expert, who is in charge of scouting relevant data sources. A *Reputation Monitor* evaluates the services from the catalogue, and assigns measures based on objective reputation criteria (e.g., institutional reputation). Such assessment is continuously updated. The output of the assessment activity is a *Reputation Descriptor* storing values for the different reputation measures.

The *Broker* selects the services that best match the specified user's settings (e.g., information needs expressed in the queries that the user submits to the platform) and

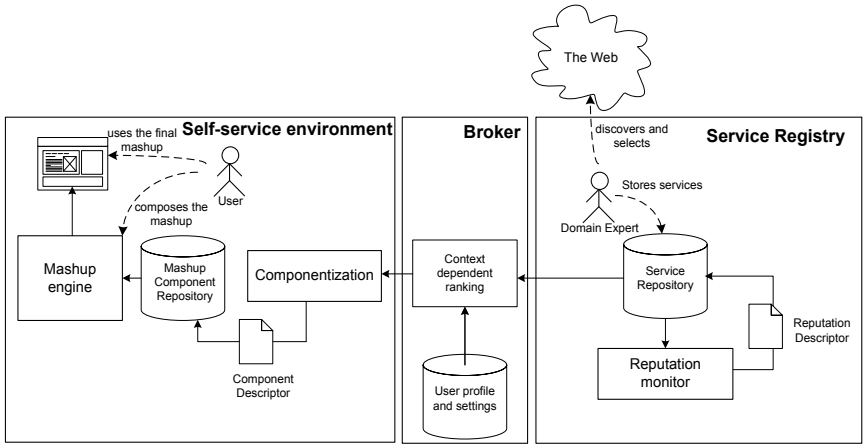

Figure. 1- INTEREST architecture.

the user profile. The user profile can represent preferences about the services to be used as data sources (e.g., Lonely Planet vs. TripAdvisor), or also preferences about a specific domain (e.g., the preferred hotel category or the preferred food). The services selected by the broker are then used by the self-service environment that helps users create their personal information access environment.

The *Self-Service Environment* is centred on the availability of a *mashup engine*, through which the user can select some relevant components from the *Mashup Component Repository* and combine them to generate new value. The components available in the repository are those previously selected by the broker. In order to be combined into a mashup, these services need to be *componentized*, i.e., each service must be associated with a descriptor highlighting the properties which are useful for combination and choreography purposes. When the services selected by the broker are

not provided with a proper user interface, as it happens for pure data services, the componentization process also requires the generation of a presentation layer. This process implies the selection of some visualization widgets (from a widget repository) that best match the operations and the exchanged data, as indicated by the service descriptors (WSDL descriptor, API, or also additional service profiles).

3 Preliminary Empirical Results

This section presents some preliminary results of the reputation assessment, which are related to the *sentiment analysis*. This is the first step performed by the *Reputation Monitor* (Figure 1), which consists in evaluating the feeling about a given entity of interest, on the basis of the opinions released by users on specialized Web sites. The Reputation Monitor then exploits the sentiment indicator, as a benchmark to weight the source trustworthiness.

The sentiment is evaluated through a prototype sentiment analyzer, which integrates *WordNet* (Fellbaum, 1998) with *Freeling* (Atserias et al. 2006) and *SentiWordNet* (Esuli & Sebastiani, 2006). The current version of the tool has the following characteristics:

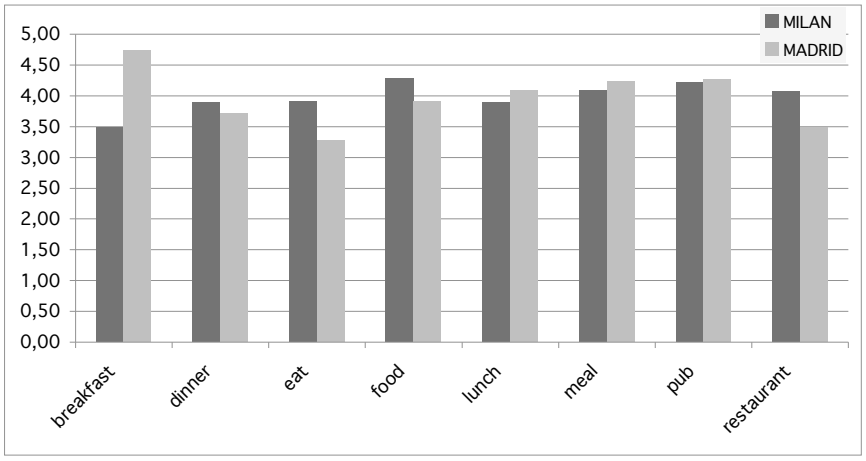

Figure. 2. A comparison between the food in Madrid and Milan.

- It focuses on *English*.
- It evaluates the sentiment of selected sources (namely Twitter, TripAdvisor, and Lonely Planet) on a single label (i.e., it accepts single-label queries).
- It extracts from the Web sources the sentences including the target label and does not consider the context of those sentences (i.e. the text in which they are embedded).

- It includes a set of rules to process sequences of adjectives and adverbs (e.g. “not that interesting” or “very interesting”).
- It selects the correct meaning of words within their synset (set of synonyms inside WordNet) based on the concept of domain. In particular, the tool focuses on the tourism domain.

Figure 2 reports the results of a comparison between Madrid and Milan on the quality of their food. The following labels have been considered: food, dinner, lunch, meal, breakfast, restaurant, pub, eat. A total of 250 sentences have been analyzed for each city. The graph reports an aggregate evaluation of sentiment obtained as a simple average of the sentiment assessed for each label. Sentiment is rated on a 1 to 5 scale, where 1 is “very negative” and 5 is “very positive”. In this dataset we obtained medium quality accuracy, with a precision of 66% and a recall of 64% on a 5-level based evaluation. Such values get noticeably better in the case of a 3-level based evaluation, obtaining a precision of 95% and a recall of 92%. Precision and recall have been assessed through a manual evaluation of the tool's results on the sample input described above.

4 Discussion and Future Work

This paper has presented some preliminary results about the development of a Web environment where users can access, select and compose reputation services. Such results especially focus on the automatic analysis of reputation, which is an important enabler of a number of services for both companies and their customers. Given the considerable effort that companies allocate to marketing analyses, the automated evaluation of reputation represents a potential breakthrough innovation. Finding the comments that most negatively impact on the company’s reputation is important in order to plan, define, and enact targeted response actions, or even compare their products or services with those offered by competitors. In some domains, such as music or theatre, these response actions can represent a key component of the cultural mission of a number of no-profit organizations. Firms operating in the tourism industry, such as companies organizing trips, hotels, but also public bodies, can get multiple advantages as well, being interested in monitoring their online reputation by querying the most dependable and authoritative sources.

We finally believe that reputation assessment can be beneficial for the end-users, especially in those contexts where users must take decisions based on other people’s opinions. We conducted an online survey using some popular social networks, asking people whether they considered reputation services important. Results confirmed our perception. In four days we received 90 replies from all over Europe and 82% of the interviewees judged the system very useful.

Our current efforts are devoted to refining the algorithm for the reputation assessment, for example by introducing term clustering to improve the analysis. We are also developing the visual environment to support the users: *i)* to personalize the reputation services by means of a suitable presentation layer, and *ii)* to compose the

so obtained components into mashup applications able to highlight relevant relationships among the results of different services.

References

Atserias, J., Casas, B., Comelles, E., González, M., Padró, L. & Padró, M. (2006). FreeLing 1.3. Syntactic and semantic services in an open-source NLP library. Proc. of the 5th Int. Conf. on Language Resources and Evaluation (LREC'06). Genoa, Italy.

Barbagallo, D., Cappiello, C., Francalanci, C. & Matera, M. (2009). Reputation Based Self-Service Environments. Proc. of the ICWE Workshop ComposableWeb'09: 12-17. San Sebastian, Spain.

Brin, S. & Page, L. (1998). The anatomy of a large-scale hypertextual Web search engine. Proc. of WWW99. Elsevier Science Publishers B. V., Amsterdam, The Netherlands, 107-117.

Esuli, A. & Sebastiani, F. (2006). SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. Proc. of the 5th Int. Conf. on Language Resources and Evaluation (LREC'06): 417-422, Genoa, Italy.

Fellbaum, C. (ed.) (1998). WordNet: An Electronic Lexical Database. MIT Press.

Gil, Y. & Artz, D. (2006). Towards content trust of Web resources. Proc. of WWW06. ACM, New York, NY, 565-574.

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *J. ACM* 46, 5 (Sep. 1999), 604-632.

Malik, Z. & Bouguettaya, A. (2009). RATEWeb: Reputation Assessment for Trust Establishment among Web services. *VLDB Journal* 18, 4 (Aug. 2009), 885-911.

Medjahed, B. & Bouguettaya, A. (2005). Customized Delivery of E-Government Web Services. *IEEE Intelligent Systems* 20, 6 (Nov. 2005), 77-84.

Murugesan, S. (2007). Understanding Web 2.0. *IT Professional* 9(4): 34-41.

Skopik, F., Truong, H. & Dustdar, S. (2009). Trust and Reputation Mining in Professional Virtual Communities. Proc. of ICWE 2009. Lecture Notes In Computer Science, vol. 5648. Springer-Verlag, Berlin, Heidelberg, 76-90.

Yu, J., Benatallah, B., Saint-Paul, R., Casati, F., Daniel, F. & Matera, M. (2007). A framework for rapid integration of presentation components. Proc. of WWW07: 923-932.