Web Intelligence and Communities

Pierre Maret
Hubert Curien Laboratory
Université Jean Monnet
Saint-Étienne, France

pierre.maret@univ-st-etienne.fr

Rajendra Akerkar
Vestlandsforsking
Sogndal
Norway

rak@vestforsk.no

Laurent Vercouter
LITIS Laboratory
INSA de Rouen
France

laurent.vercouter@insa-rouen.fr

ABSTRACT
The World Wide Web (WWW) provides precious means for communication, which goes far beyond the traditional communication media. Web-based communities have become imperative spaces for individuals to seek and share expertise. Networks in these communities usually differ in their topology from other networks such as the World Wide Web. In this paper, we explore some research issues of web intelligence and communities. We will also introduce the WI&C'15 workshop's goal and structure.

Categories and Subject Descriptors

I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence

Keywords
Web intelligence, web communities, ICT, knowledge, social networks, intelligent agents.

1.   INTRODUCTION

In recent years, web communities received a visible level of attention from the research community in many disciplines [2]. The concept of (artificial) intelligence has been broadly debated from numerous aspects. The web intelligence has been recognized as a way to investigate the underlying role as well as practical impact of Artificial Intelligence (e.g. knowledge representation, planning, knowledge discovery and data mining, intelligent agents, and social network intelligence) and advanced Information Technology (e.g. wireless networks, ubiquitous devices, social networks, and wisdom web) on the next generation of web-empowered products, systems, services, and activities.

Web communities [3] use networked technology, mainly the Internet, to establish alliance across geographical obstacles and time zones. In contrast to conventional communities, web communities in World Wide Web differ in several respects. Conventional communities are place-based and have membership according to norms. Crowd dynamics frequently dominate individual expression. There is a definite edge between membership and differentiation. For instance, it is clearly stated who is a member and who is not. Whereas, web communities appear according to identification to an idea or task. They are

organized around an activity, and they are formed as a need arises. Interestingly web communities do not need formal boundaries because the members cannot see each other, rules do not dominate as much as in conventional communities, thus letting superior individual control. In other words, the WWW, becomes the ''place'' for the community; thus networked communication has augmented the parameters of what is called as a community.

In technical term, a web community is an architecture made up of a set of actors and a web of linkages between these actors. The web community perspective provides a clear way of analysing the structure of whole community entities and identifying essential links among various nodes of the network. The study of these structures uses intelligent analysis to identify local and global patterns, locate influential entities, and examine complex network dynamics.

2.   SOME RESEARCH ISSUES
One of the key objectives of web intelligence and communities research is analysis of the patterns of individual/community behaviours and community dynamics.

So, it is vital to explore the dynamics of web communities, for example, the variations and interactions between web communities and homophily (which the tendency of individuals to associate and bond with similar others). While links in the former are explicit links between individuals (e.g., by the is-friend-of relation), links in the latter result from some similarity in behaviour (e. g., has-tagged-the-same resource-as). To explore how homophily can initiate social relations, and in which situations such a correlation can be perceived.

Another area is characteristics of semantic spaces as the standard model of meaning representation in quantitative semantics. We argue that these characteristics are challenged from the perspective of web communities and the potentials which they offer in terms of exploring semantic and pragmatic data.

Gathering and analysis of enormous amount of data are foundational challenges in every area of science and engineering. In a situation, where every distinct member of a Web community can easily contribute to the data tsunami, data may differ in terms of subjectivity and significance, varying from personal opinions and estimations to largely established practices and well-documented scientific results. This promotes the utilization of the synergy between human and machine reasoning when designing systems to support such processes within Web communities. Development of data mining know-hows for pattern and dependencies discovery within big datasets is indeed of huge benefit [1]. But, in spite of advancement made in the field of computational analysis, there are many patterns that humans can

easily detect but computer algorithms struggle to estimate. Interpretation of analysis' results is a challenging issue here, in that getting results from the execution of a knowledge discovery algorithm is hardly enough; extra information is needed concerning how each result came out and based on which input.

In some cases, web communities are not explicitly or publicly represented, especially when web technologies are used to enable people sharing a common objective to exchange or propagate information without being recognized. Detection and recognition of such communities raise important scientific challenges. Data analysis technics are required to recognize typical features. But these features have to be compared to identified individual and collective behavioural patterns that have to be built. New advances in the fields of user modelling, dynamic graph analysis and multi-agent systems will allow to identify the activity of these hidden web communities and to recognize it.

There is no web community without social beings. Hence, we have to take into consideration human's perception of objects of the real world and manner of their description for web communities. We need to focus on two vital issues: How we can make our web community more personal, i.e. different for different users? , and how we can make our web community more optimal, i.e. more passable as a model of a real world we would like to operate in?

3. PAPERS IN THE WORKSHOP

The papers selected for this workshop deal with some of the significant issues in the field.

Social Networks and the Semantic Web: a retrospective of the past 10 years is the title of the keynote speech. This talk presents to domain experts and newcomers an overview of existing bond between social networks and Semantic Web techniques and points out future research directions.

In WI&C'15 workshop, three papers are selected for presentation. These papers are:

An evaluation of SimRank and Personalized PageRank to build a recommender system for the Web of Data.

This paper evaluates SimRank and personalized PageRank for calculating similarity on RDF graphs. The algorithms exploit structural context, and they can be utilized on the Linked Data graphs for improving content-based recommendation systems. The experiments focus on recommendations as a use-case and for evaluating the approach.

Exploriometer: leveraging personality traits for coverage and diversity aware recommendations.

The paper describes an Exploriometer, which is a new metric that offers a personality trait for users based on their rating behavior. The metric has been evaluated by experiments with two real-world datasets (MovieLens and Yahoo! Music). The experimental evaluation of the technique for enhanced neighbor selection is reported.

A bidimensional user profile to discover unpopular Web sources.

The paper proposes a personalized focused crawler to help users discovering relevant information sources in the context of Experts in Intelligence Analysis (EIA). The purpose is to cover the thematic of the information needs as well as specific and accurate terms. It further states a new similarity measure that is used in a Web source discovery system called DOWSER. DOWSER aims at providing users with new sources of information related to their needs without considering the popularity of a page.

4. REFERENCES

[1] Akerkar, R. 2013. *Big Data Computing*. Taylor and Francis/CRC.

[2] El Morr, C. and Maret P. 2012. *Virtual Community Building and the Information Society: Current and Future Di*rections. IGI Global.

[3] Rheingold, H. 2000. *The Virtual Community: Homesteading on the Electronic Frontier*. MIT Press, London.