Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 50 (2015) 524 – 530

Procedia Computer Science

2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)

Towards the development of a framework for socially responsible software by analyzing social media big data on cloud through ontological engineering

Alok Chauhan, V. Vijayakumar, Rajiv Vincent, Pradeep K V
School of Computing Sciences & Engineering, VIT Chennai – 600127, Tamil Nadu, India

Abstract

A socially responsible internet is the need of the hour considering its huge potential and role in educating and transforming the society. Social computing is emerging as an important area as far as development of next generation web is concerned. With the proliferation of social networking applications, vast amount of data is available on cloud, which may be analyzed to gain useful insight into behavioral and linguistic patterns of different cultural and socio-economic groups further classified on the basis of gender and age etc. The idea is to come up with an appropriate framework for socially responsible software artifacts. These artifacts will monitor online social network data and analyze it from the perspective of socially responsible behavior based on ontological engineering concepts. Identification of socially responsible agents is such an example, though based on a different approach. More examples may be taken from literature dealing with microblog analytics, social semantic web, upper ontology for social web, and social-network-sourced big data analytics. In the present work, it is proposed to focus on analysis/monitoring of socially responsible behavior of social media big data and develop an upper level ontology as the framework/tool for such an analytics.

Peer-review under responsibility of scientific committee of 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)

Keywords: *Social computing, socially responsible software, ontological engineering, upper level ontology, madhyasth-darshan.*

1. Introduction

The concept of socially responsible software may be applied in diverse fields such as artificial intelligence, software architecture, distributed computing, semantic web etc. Its full potential may be realized with the arrival of new computing paradigms such as service-oriented architecture. Performance analysis of responsible agents in the

Peer-review under responsibility of scientific committee of 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15)
doi:10.1016/j.procs.2015.04.026

framework of multi-agent system design has given encouraging results [1]. The idea is to introduce a new computer level known as the social level [1]. It is proposed to extend it to social semantic web. There is exponential growth in the amount of information on the web. Semantic web deals with the issue of efficiently managing this information. At present, a bottom-up approach is suggested for the evolution of social semantic web [2]. But this may not be a suitable approach for a socially responsible semantic web. A top-down approach based on a socially responsible upper level ontology for social semantic web may provide us a solution. The proposed methodology suggests meta language analytics of social media big data on cloud as a verification mechanism for our model.

2. Related Work

The idea of socially responsible software goes as far back as its reference in multi-agent system design. With the proliferation of social web, which is the dominant feature of Web 2.0, the idea of socially responsible software deserves a new meaning. The survey is categorized as follows:

2.1 Socially Responsible Agents

Software agents are considered to be an important technology for solving complex problems. These agents are autonomous software components, which follow distributed computing paradigm. Collaboration among agents is the key requirement. That way, they are also examples of upcoming co-operative computing. Looking at the development in this area, in the beginning more emphasis was put on the overall system design aspect as compared to degree of autonomy of an individual agent. This approach is known as reductionist approach. Later developments saw emergence of a new approach known as constructionist approach in which individual agents are empowered more while less emphasis is given to holistic view of the system. Both the approaches have their advantages as well as disadvantages. The concept of socially responsible agent should be seen in terms of its problem solving behavior wherein each agent being an autonomous agent is not only responsible for its individual problem solving behavior but since it is also part of a system/community of agents its social responsibility is expressed through its role in providing cooperation to other agents. Hence, what is required is a balance in individual/community roles [1]. The architecture of an agent as shown in Fig. 1 [1] clarifies it further.

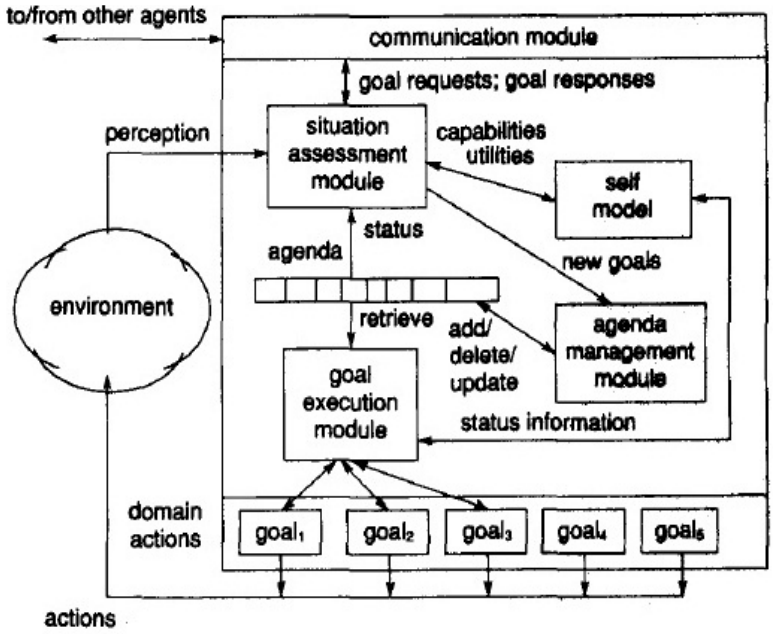

Fig. 1: Agent architecture

2.2 Microblog Analytics

Social media big data analytics on cloud is going to be the next real challenge for computer scientists. Huge amount of data is being generated on daily basis by microblog service like Twitter. People are interested in mining data relevant to them and filtering out noise component. Techniques employed include creation of various types of filters and classifiers even dynamically, monitoring of environment and ability to support supervision. It also involves stream processing and big data analytics on cloud [2]. Fig. 2 [2] explains it further.

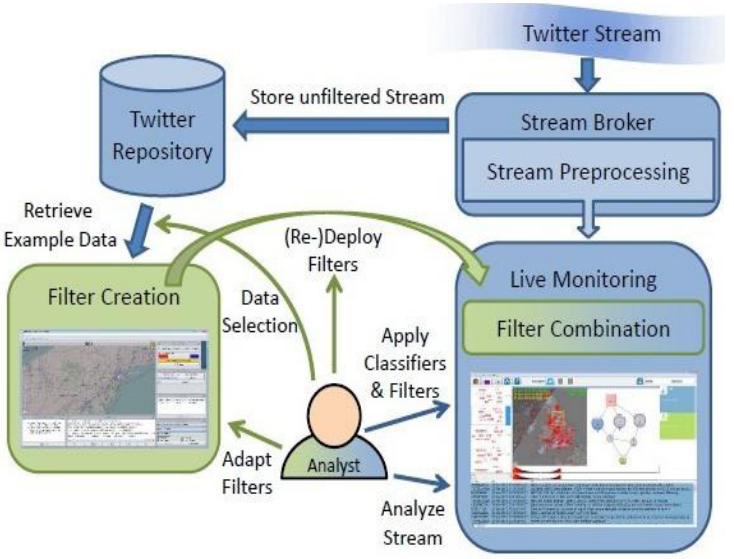

Fig. 2: Training of classifiers as well as creation of statistically motivated, weighted keyword filters.

2.3 Social Semantic Web

Emergence of social web sites is a major trend making Internet more democratic medium where people cannot only share their knowledge but also be able to do some collaborative work as well. This creation of knowledge by people also known as crowdsourcing is going to affect the way Internet will be utilized in future. However, this trend throws up some new challenges. Non-availability of any common standard for this shared knowledge does not allow us to realize its full potential. Semantic web community is trying to find solution to this mega-challenge. Initiatives like SIOC (Semantically Interlinked Online Communities) and FOAF (Friend-of-a-Friend) are examples of such efforts which may make social web interoperable. SIOC Core Ontology gives us a framework to manage semantic issues involved with online communities. At present, mostly these communities operate in standalone mode which hampers reuse of precious information. SIOC Core Ontology provides us with necessary artifacts to handle blogs, wikis, podcasts and online social networks data. SIOC uses RDF (Resource Description Framework) as a metadata model [4]. Another approach to social semantic web suggests bottom-up technique to generate ontologies from folksonomies. It premises on the fact that crowdsourcing is better than input from few experts. It puts emphasis on simplicity and efficiency [3]. Social semantic web may be illustrated as shown in Fig. 3 [4].

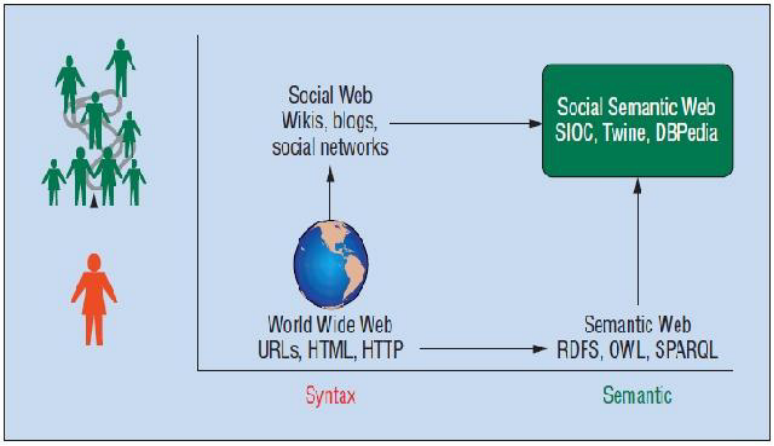

Fig. 3: The social semantic web

2.4 Upper Ontology

FOAF has been used to represent social relationships in various ontologies. But efforts have been made to model social contexts such as social relations and institutional relations. It can be further extended to include social archetypes as well [5].

2.5 Big Data Analytics

Exponential growth in information generated by organizations, online communities and sensor based systems is causing a phenomenon known as data deluge. Big data analytics shows hidden patterns in accumulated data which generates new knowledge benefitting the organizations and society as a whole. There is much interest in big data originating from social networks [6]. One can see this dynamics in Fig. 4 [6].

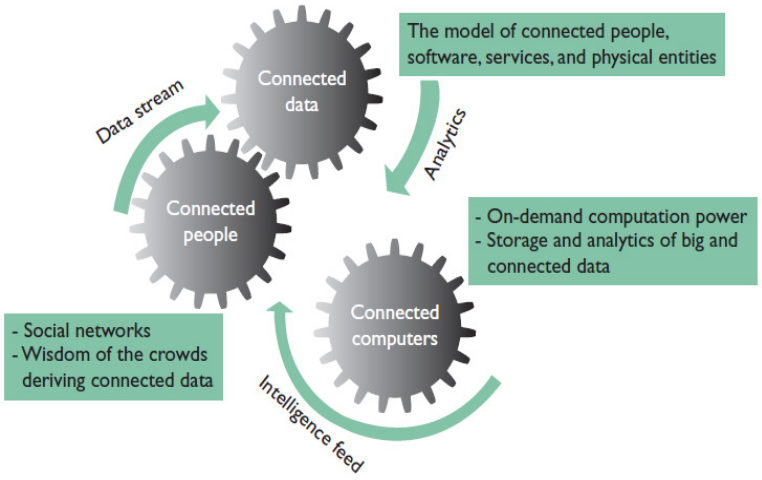

Fig. 4: The virtuous circle.

Big data is characterized by three Vs known as volume, velocity and variety. Its analysis involves concepts from complex network theory. New parallel programming frameworks such as MapReduce is used. NoSQL systems are gaining priority. Elasticity of such systems help. Stream processing is another area of importance. Security issues are another big challenge. Aggregate data analytics widens the scope across various domains [6].

3. OBSERVATIONS

It is proposed to look at the problem at hand from the perspective of ontological engineering. Ontological engineering encompasses following issues:

3.1 Building Ontologies

Ontology development process involves management, development oriented, and support activities. The process does not identify the order of these activities. Ontology life cycle identifies the order [7].

3.2 Evaluating Ontologies

Ontology evaluation involves verification, validation, assessment, consistency, completeness and conciseness. While evaluating taxonomy, errors may be found due to inconsistency, incompleteness and redundancy. Inconsistency errors may be further classified as circularity, partition and semantic errors [7].

3.3 Implementation Using Traditional Ontology Languages

Ontologies may be implemented in various ontology languages available. Proper selection of an ontology language is very important. These languages differ a lot in terms of their underlying knowledge representation paradigms, which are as diverse as frames, description logic, first (and second) order logic, semantic networks etc [7].

3.4 Ontology Markup Languages

These languages act as foundation for semantic web. Lot of development is still taking place in this field so things may change in future. These languages are based on XML [7].

3.5 Ontology Tools

Ontology tools may be classified as development tools, evaluation tools, merge and alignment tools, annotation tools, querying tools, inference engines, and learning tools [7].

4. Proposed SCHEME

The proposed four stage scheme is explained in the form of a diagram as shown below in Fig. 5.

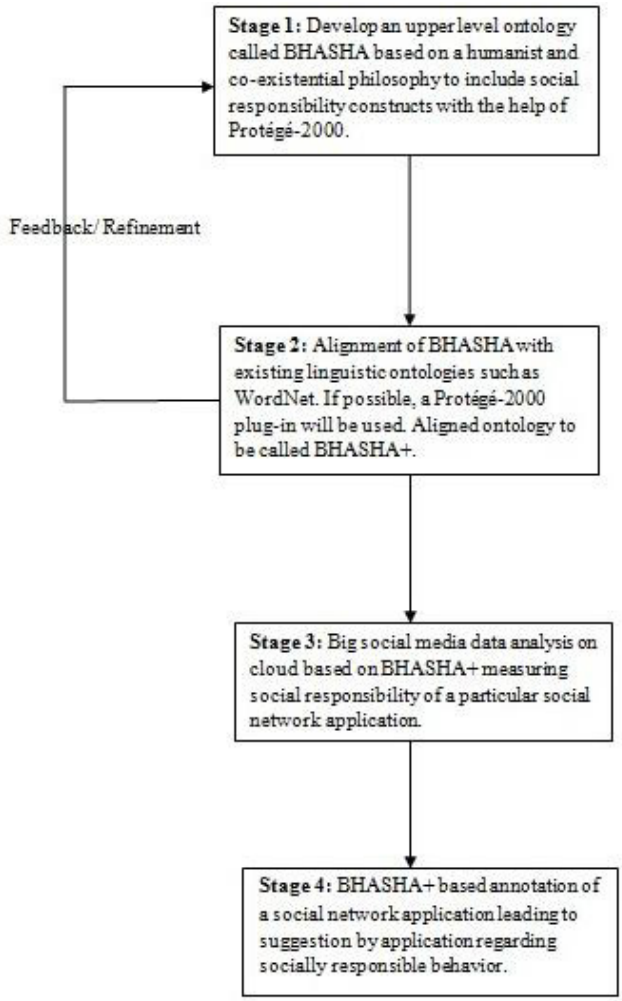

Fig. 5: Socially responsible ontological engineering.

Stage 1 is about development of a new upper level ontology (BHASHA), focusing on societal aspects of human behavior and derived from a newly proposed humanist and co-existential philosophy called madhyasth–darshan [8].

Stage 2 will deal with alignment of proposed ontology with some popular existing linguistic ontologies such as WordNet through an iterative process (incremental development). Aligned ontology is called BHASHA+.

Stage 3 will involve ontological analysis of social media big data using BHASHA+ as a classifier/ filter.

Finally, stage 4 is about BHASHA+ based annotation of a social web application to monitor in real time socially responsible content/behavior of users and positive suggestions from the application in this regard as well as a test for the proof-of-concept.

5. CONCLUSION AND FUTURE WORK

The definition of socially responsible software is at present limited to developing a socially responsible agent in the abstract concept of problem solving process. This paper has extended this concept to a real-life situation in terms of behavior of a social network application. A broad framework for the development of such applications has been presented based on theory and practice of ontological engineering.

Apart from implementing the proposed scheme, the concept will be taken even further to include some other domain ontologies also in the framework apart from linguistic ontologies.

References

[1] N.R. Jennings, J.R. Campos, 'Towards a social level charactreisation of socially responsible agents," IEE Proc. – Softw. Eng., Vol. 144, No. 1, 1997, pp. 11-25.

[2] H. Bosch, D Thom, F. Heimerl, E. Puttmann, S. Koch, R. Kruger, M. Worner, T. Etrl, 'ScatterBlogs2: Real- Time Monitoring of Microblog Messages Through User-Guided Filtering," IEEE transactions on visualization and computer graphics, vol. 19, no. 12, December 2013.

[3] A. Mikroyannidis, 'Toward a Social Semantic Web," IEEE Computer, Nov. 2007, pp. 113-115.

[4] U. Bojars, J.G. Breslin, V. Peristeras, G. Tummarello, S. Decker, 'Interlinking the Social Web with Semantics," IEEE Computer Society, May/June 2008.

[5] J. Grabarske, D. Heutelbeck, 'An Upper Ontology for the Social Web," IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining', 2012.

[6] W. Tan, M.B. Blake, I. Saleh, S. Dustdar, 'Social-Network-Sourced Big Data Analytics," IEEE Internet Computing, 2013.

[7] A. Gómez-Pérez, M. Fernández-López, O. Corcho, Ontological Engineering, Ist ed., Springer, 2011, pp. 107-362.

[8] http://madhyasth-darshan.info/