

A Semantic-Driven Adaptive Architecture for Large Scale P2P Networks

Athena Eftychiou, University of Surrey, UK

Bogdan Vrusias, University of Surrey, UK

Nick Antonopoulos, University of Derby, UK

ABSTRACT

The increasing amount of online information demands effective, scalable, and accurate mechanisms to manage and search this information. Distributed semantic-enabled architectures, which enforce semantic web technologies for resource discovery, could satisfy these requirements. In this paper, a semantic-driven adaptive architecture is presented, which improves existing resource discovery processes. The P2P network is organised in a two-layered super-peer architecture. The network formation of super-peers is a conceptual representation of the network's knowledge, shaped from the information provided by the nodes using collective intelligence methods. The authors focus on the creation of a dynamic hierarchical semantic-driven P2P topology using the network's collective intelligence. The unmanageable amounts of data are transformed into a repository of semantic knowledge, transforming the network into an ontology of conceptually related entities of information collected from the resources located by peers. Appropriate experiments have been undertaken through a case study by simulating the proposed architecture and evaluating results.

Keywords: Collective Intelligence, Distributed Information Retrieval, Domain Ontology, Peer-To-Peer (P2P) Networks, Semantic Web

INTRODUCTION

The Semantic Web idea (Berners-Lee, Hendler, & Lassila, 2001) in unstructured adaptive P2P networks is an approach to represent, manage, and retrieve distributed knowledge in an efficient manner. To pursue the Semantic Web vision, metadata is an essential add-on for information resources. This rich representation

of data aims to improve knowledge discovery and data management. Traditionally, adding metadata to resources is a manual and expensive process and probably the main cause of the slow growth and difficult implementation of the Semantic Web.

Various efforts have been made from the P2P research community to create P2P systems that can discover resources efficiently and accurately. Standard P2P searching technologies like Gnutella 2 (Stokes, 2002) produce irrelevant

DOI: 10.4018/jghpc.2010100102

search results with low precision rate. Kazaa (Liang, Kumar, & Ross, 2004) is a very popular super-peer based application, which appears to offer better scaling; however it inherits the limitations that come with flooding algorithms. Structured P2P architectures on the other hand, like Chord (Stoica, Morris, Karger, Kaashoek, & Balakrishnan, 2001), CAN (Ratnasamy, Francis, Handley, Karp, & Schenker, 2001) and Pastry (Rowstron & Druschel, 2001), guarantee retrieval of existing network resources with the use of Distributed Hash Tables (DHTs). The maintenance of a distributed index however comes with high cost, as additional traffic is generated for maintaining the routing information in a highly dynamic environment.

Trying to overcome the problems mentioned above, there are a number of P2P research systems (Cudré-Mauroux, Agarwal, & Aberer, 2007; Ehrig et al., 2003; Nakauchi, Morikawa, & Aoyama, 2004; Nejdil et al., 2002; Tatarinov & Halevy, 2004) that try to encompass the semantic technology notion, to represent meaning and knowledge, as well as to use reasoning for retrieving the knowledge. Knowledge needs to be encoded in a structured form to become widely accessible. An ontology structure (Jepsen, 2009) is an important part of the semantic web as it represents knowledge at the level of concepts; ontology provides a shared understanding through conceptualisation. In P2P systems, a domain specific ontology can be used for classifying network resources to ontology related concepts. Related research methods, are trying to exploit the benefits provided by semantics through an ontology structure in order to enhance data management and to improve resources discovery. As a result of that, the query can also be categorised to a specific concept and be routed to the peer that supports the specific concept.

An ontology-based P2P topology for service discovery has been described by (Schlosser, Sintek, Decker, & Nejdil, 2003), where the network is organised in a HyperCup topology. Information or services that peers provide are categorised to general concepts; these concepts and their relationships form the network ontol-

ogy. Peers with similar interests are organised in a concept cluster. Each cluster is assigned a combination of concepts which best describe the peers that belong to the cluster. This network organisation aims to efficiently route the query to peers that can satisfy it. This technique uses broadcasting for query forwarding; when the query reaches the most appropriate cluster is then broadcasted to all peers. The network uses globally known ontologies for clustering its nodes. This approach however is an example of structured P2P networks. Yongxiang Dou and Xiaoxian Bei (2008) also presented a semantic information retrieval system based on a hybrid ontology integration approach. In hybrid approaches, each network source employs its own local ontology but all ontologies are built upon a global shared vocabulary. The authors are focusing on the problems involved with information representation and integration in P2P semantic retrieval networks. This research and the system proposed by the authors are in preliminary stages and no experiments have proven the validity of the approach yet.

Summarising, P2P networks suffer from low quality of results, large amounts of poorly managed data and increased network traffic. The main motivation of this work is to attempt and resolve these issues by comprising the Semantic Web idea in unstructured adaptive P2P networks. The importance of this work is reflected in the challenge to pursue the Semantic Web vision by understanding the semantics of the data, to provide fault tolerance and scalability improvements. This can be achieved by adapting P2P architectures and combining the Semantic Web technologies with P2P networks, enhancing in this way the P2P performance in terms of accuracy, speed and traffic.

Consequently, the current research focuses on creating a dynamically adaptive semantic-driven P2P topology for managing and distributing knowledge in an efficient manner. The main aim is to transform the resources of a large scale P2P system from unmanageable amounts of data into a structured knowledge based repository of indexed resources. The implementation of a Collective Intelligence (CI)

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/semantic-driven-adaptive-architecture-large/47208

Related Content

Performance Evaluation of Reactive Routing in Mobile Grid Environment

L. Shrivastava, G. S. Tomar and S. S. Bhaduria (2011). *International Journal of Grid and High Performance Computing* (pp. 45-53).

www.irma-international.org/article/performance-evaluation-reactive-routing-mobile/56355

A Novel Approach of Symmetric Key Cryptography using Genetic Algorithm Implemented on GPGPU

Srinivasa K. G., Siddesh G. M., Srinidhi Hiriyannaiah, Anusha Morappanavar and Anurag Banerjee (2016). *Emerging Research Surrounding Power Consumption and Performance Issues in Utility Computing* (pp. 283-303).

www.irma-international.org/chapter/a-novel-approach-of-symmetric-key-cryptography-using-genetic-algorithm-implemented-on-gpgpu/139849

An Effective Volleyball Trajectory Estimation and Analysis Method With Embedded Graph Convolution

Guanghai Huang (2023). *International Journal of Distributed Systems and Technologies* (pp. 1-13).

www.irma-international.org/article/an-effective-volleyball-trajectory-estimation-and-analysis-method-with-embedded-graph-convolution/317936

An Introduction to Controlflow and Dataflow Supercomputing

Miloš Kotlar (2021). *Handbook of Research on Methodologies and Applications of Supercomputing* (pp. 1-4).

www.irma-international.org/chapter/an-introduction-to-controlflow-and-dataflow-supercomputing/273391

From Enabling to Ensuring Grid Workflows

Junwei Cao, Fan Zhang, Ke Xu and Lianchen Liu (2009). *Quantitative Quality of Service for Grid Computing: Applications for Heterogeneity, Large-Scale Distribution, and Dynamic Environments* (pp. 46-73).

www.irma-international.org/chapter/enabling-ensuring-grid-workflows/28270