Swarm and Evolutionary Computation 29 (2016) 73–93

Contents lists available at ScienceDirect

Swarm and Evolutionary Computation

journal homepage: www.elsevier.com/locate/swevo

Regular Paper

Bio-inspired search algorithms for unstructured P2P overlay networks

Vesna Šešum-Čavić *, Eva Kühn, Daniel Kanev

Vienna University of Technology, Institute of Computer Languages, Argentinierstr. 8, 1040 Vienna, Austria

ARTICLE INFO

Article history:
Received 22 August 2015
Received in revised form
9 December 2015
Accepted 15 March 2016
Available online 4 April 2016

Keywords:
Bio-inspired intelligence
Slime molds
Unstructured P2P overlay networks
Intelligent lookup
Location and retrieval of information

ABSTRACT

Efficient location and manipulation of complex and often incomplete data is a difficult, challenging task in nowadays extremely complex IT systems and on the Internet, overwhelmed with a huge amount of information. The problem itself is present in numerous different practical use-cases (e.g., in P2P streaming applications that rapidly gain more attention) and refers to the selection of the proper, efficient search algorithm. Research and commercial efforts resulted in a prolific offer of different algorithms that try to address this problem in the best possible way. Due to the huge complexity, intelligent algorithms are the most promising ones. However, everyday changing conditions impose finding even more advantageous approaches that will better cope with the problem, or at least address some "corner cases" better, than previously realized ones. In this paper, we propose a self-organizing approach inspired by bio-intelligence of slime molds that possesses distributive and autonomous properties with the goal to achieve a good query capability. A slime mold mechanism is adapted for search in an unstructured P2P system, and compared with Antnet and Gnutella search mechanisms. The benchmarks cover parameter sensitivity analysis, and comparative analysis. To validate the obtained results, a statistical analysis is performed. The obtained results show good scalability of slime mold algorithm and point to the selected "corner" cases where the slime mold algorithm has a total good performance (measured by different metrics).

1. Introduction

Bio-inspired metaheuristics are important and inevitable for the optimization and robustness of highly dynamic distributed systems where autonomous agents interact without a central control. The behavior of such complex systems is typically unpredictable. A very useful concept in the adaptation of complex systems, in general, is the one of self-organization. There is a constant necessity for self-organizing mechanisms in distributed systems [19]. Bio-inspired intelligence could help highly dynamic systems to cope with environmental changes by providing some properties inherited from biological systems: every item in the population makes local decisions, and behaves and acts in a decentralized manner.

One of nowadays constantly arising problems is the efficient *location and retrieval of information* in the Internet. It becomes a more and more complex and difficult task that faces many challenges and copes with a highly dynamic nature of the Internet. An additional challenge is the manipulation of complex, and particularly, incomplete data (their efficient storing, querying and processing) imposed by an increasing complexity of systems of real-life applications. This requires an advanced approach that is able to manage and solve the above-mentioned problems in an autonomous, intelligent manner and that is sufficiently adaptable. In order to better explain the meaning of complex, incomplete data, the following example is given: Let us suppose that we want to find the name of the song heard recently and we know only some part of its title. We know neither who performs it, nor the complete title, nor which year the song originates from.

In order to properly address such a class of problems, an efficient search mechanism in the "underlying" unstructured P2P overlay network should be found. More details about why unstructured P2P overlay networks are in the focus of this paper can be found in Section 2.2. Generally, unstructured P2P overlay networks support very well dynamics and complex queries, but the disadvantage is that they do not scale well, which is the starting point for an improvement. Also, the path-optimization process in search should be improved and "reinforced" in order not only to compute the shortest path per se, but the optimal path for the amount of resources involved. Therefore, the main research question we start with concerns the efficacy of an intelligent search in unstructured P2P overlay networks.

In this paper, we propose the usage of an intelligent lookup mechanism based on the lifecycle of slime molds in fully unstructured P2P overlay networks. The proposed solution is a self-

* Corresponding author.
E-mail addresses: vesna@complang.tuwien.ac.at (V. Šešum-Čavić), eva@complang.tuwien.ac.at (E. Kühn), dkanev@complang.tuwien.ac.at (D. Kanev).

http://dx.doi.org/10.1016/j.swevo.2016.03.002
2210-6502/

organizing one obtained as a combination of unstructured P2P and space based computing [24,25] for searching and retrieving data concurrently. The subjects in our architecture are software agents that perform the roles of artificial species. Note that the software itself does not solve the lookup problem, but serves as a necessary basement for the used algorithms and abstracts the general requirements. It allows exchanging different, bio-inspired as well as other algorithms simply through "plugging", and supports many different network topology settings through configuration. Therefore, one swarm intelligence based search (AntNet) and one conventional search (Gnutella) are plugged-in and compared to the intelligent lookup based on the slime mold behavior. As the main point in this paper concerns the search algorithm, the main characteristics of underlying software architecture will be only very briefly mentioned in Section 1.1 without full description.

Slime mold intelligence is a novel approach for the intelligent lookup mechanism, because the slime molds do not compute the shortest path *per se*, but the optimal path for the amount of resources involved [2]. Therefore, our expectations were to improve already existing results, i.e., to locate cases in which slime mold based search mechanism gives a benefit.

The *novelty* and *contribution* of this paper include:

- an adaptation of slime mold intelligence for lookup in unstructured P2P networks: it is for the first time used for this type of search;
- a "fair" comparison between selected intelligent and non-intelligent algorithms;
- evaluation of scalability of slime mold algorithm and identification of "corner" cases where the slime mold algorithm has a total good performance (i.e., measured by different metrics); and
- an adequate support of manipulation with "incomplete" queries (where some information is missing).

Our motivation came from the specific use-case of P2P video-streaming. Video-over-IP applications have recently attracted a significant attention and a large number of users on the Internet. Several P2P streaming systems have been deployed to provide live and on-demand video streaming services on the Internet at low server cost and with minimal dedicated infrastructure [13,47]. However, the basis of the search success is the adequate algorithm, which is the main point of this paper.

1.1. Problem identification in the real-world

P2P applications have recently gained a huge success in the area of Voice-over-IP (VOIP) and instant messaging. The benefits of the P2P technology in VOIP are indisputable [28]. If a client/server model is used and a single server is dedicated to handle the multimedia traffic, soon the limits of the server will be reached. A more scalable, efficient and cheap approach is to utilize the network bandwidth of different users interconnected in a global P2P network. Skype is a VOIP and instant messaging application that uses the super peer system architecture [30]. Another area, similar to VOIP, is P2PTV, whose applications distribute live video streams, mostly TV channels [44]. They share similar principles and architectures as the VOIP applications. On-demand P2P streaming applications like Spotify [23] provide the ability to consume digital content instantly and to change the playback position any time.

The underline software architecture of this current research attempts also to solve the problem of creating a P2P decentralized unstructured flat overlay network with streaming content delivery method and user collaboration. The product – the P2PStreamer-is a peer-to-peer application, which provides users the ability to share video files (encoded in H.264 format, as it is one of the modern formats used to encode video), search videos in the network using a fully distributed search algorithm, stream videos using mesh-based fully distributed Video On-Demand streaming scheme, and collaborate by sharing video comments below each streamed video. The P2PStreamer provides a security mechanism, which is centralized and implemented using the XVSM[1] P2P middleware [25]. The security mechanism prevents usage of the system by unauthenticated users. Also, denying access to specific users for specific videos is possible. Each video in the P2P network has meta-data associated with it, and for simplicity and illustrative purpose, an implementation with three parameters will be provided: name, director, and year. Access denial is dynamic, i.e. it is evaluated based on dynamic data. Our solution takes advantage of unstructured P2P networks, space based computing and swarm intelligence. The combination of these three resources is as follows: the searching and retrieving in our unstructured P2P overlay network is realized by using bio-inspired intelligence, whereas space based computing is used for the implementation of (sub) spaces, so-called containers, in the overlay network.

The paper is structured as follows: Section 2 contains the background and related work. Section 3 describes the slime mold algorithm for the lookup in unstructured P2P networks. Section 4 contains experimental results and a thorough discussion. The conclusion is in Section 5.

2. Background

This section is organized in three subsections and contains:

(1) the outline of basic slime molds in nature (Section 2.1); the lifecycle of one type of slime molds is described; further, current applications of slime mold mechanisms are briefly mentioned;
(2) a short overview of P2P and the explanation of why *unstructured* P2P overlay networks are taken into consideration (Section 2.2), and
(3) the state-of-the-art of the intelligent search in unstructured P2P overlay networks and their classification (Section 2.3); consequently, the basic characteristics of metaheuristics mentioned in the state-of-the-art are comparatively presented, i.e., systematized in a table; at the end of the detailed analysis in Section 2.3, a kind of summary about the current related work is presented.

2.1. Slime mold in nature

Self-organization in nature and intelligent bio-mechanisms that cope with high complexity serve as an inspiration for solving different IT problems (e.g., [9,34]). Self-organization often refers to diverse pattern formation processes in the physical and biological world [10]. One intelligent bio-mechanism is the behavior of slime molds. This section describes the lifecycle of one type of slime molds with all its stages. At the end of this subsection, the current applications of slime mold mechanisms in different scenarios are listed.

The slime molds are organisms which use spores to reproduce. There are various types of slime molds. According to the modern classification [21], they can be broadly divided into amoebozoan and non-amoebozoan. Inside these groups, further sub-classification is made. This paper is focused on the one of amoebozoan slime mold, i.e., the lifecycle of the cellular slime mold: *Dictyostelium discoideum* (Dd). The Dd's lifecycle goes through five stages:

[1] http://www.xvsm.org