The BISON Project

Alberto Montresor * Ozälp Babaoglŭ *

Introduction

Modern distributed information systems are gaining an increasing importance in our every day lives. As access to networked applications become omnipresent through PC's, hand-held and wireless devices, more and more economical, social and cultural transactions are becoming dependent on the reliability and availability of distributed applications.

As a consequence of the increasing demands placed by users upon networked environments, the complexity of modern distributed systems has reached a level that puts them beyond our ability to deploy, manage and keep functioning correctly through traditional techniques. Part of the problem is due to the sheer size that these systems may reach, with millions of users and interconnected devices. The other aspect of the problem is due to the extremely complex interactions that may result among components even when their numbers are modest. Our current understanding of these systems is such that minor perturbations (e.g., a software upgrade, a failure) in some remote corner of the system will often have unforeseen, and at times catastrophic, global repercussions. In addition to being fragile, many situations (e.g., adding/removing components) arising from their highly dynamic environments require manual intervention to keep information systems functioning.

In order to deal with the scale and dynamism that characterize modern distributed systems, we believe that a paradigm shift is required that includes self-organization, adaptation and resilience as intrinsic properties rather than as afterthought. For this reason, we have started BISON (*Biology-Inspired techniques for Self Organization in dynamic Networks*), an international project partially funded by the European Commission. BISON draws inspiration from natural and biological processes to devise appropriate techniques and tools to achieve the proposed paradigm shift and to enable the construction of robust and self-organizing distributed systems for deployment in highly dynamic modern network environments.

*Department of Computer Science, University of Bologna, Mura Anteo Zamboni 7, 40127 Bologna (Italy), Email: {babaoglu,montresor}@CS.UniBO.IT

Overview of the BISON Project

Nature and biology have been a rich source of inspiration for computer scientists. *Genetic algorithms* [4], *neural networks* [7] and *simulated annealing* [8] are examples where different natural processes have been mimicked algorithmically to successfully solve important practical problems. More recently, *social insects* [2] have been added to this list, inspiring biomimetic algorithms to solve combinatorial optimization problems. *Immune networks* [9] represent a recent frontier in this area.

These processes are examples of *complex adaptive systems (CAS)* [6] that arise in a variety of biological, social and economical phenomena. In the CAS framework, a system consists of large numbers of *autonomous agents* that individually have very simple behavior and that interact with each other in very simple ways. CAS are characterized by total lack of centralized coordination. Despite the simplicity of their components, CAS typically exhibit what is called *emergent behavior* that is surprisingly complex [6]. Furthermore, the collective behavior of a well-tuned CAS is highly adaptive to changing environmental conditions or unforeseen scenarios, is robust to deviant behavior (failures) and is self-organizing towards desirable configurations.

Parallels between CAS and advanced information systems are immediate. BISON will exploit this fact to explore the possibility of using ideas and techniques derived from CAS to enable the construction of robust, self-organizing and self-repairing information systems as ensembles of autonomous agents that mimic the behavior of some natural or biological process. In our opinion, the application of CAS will enable developers to meet the challenges arising in dynamic network settings and to obtain global properties like resilience, scalability and adaptability, without explicitly programming them into the individual agents. This represents a radical shift from traditional algorithmic techniques to that of obtaining the desired system properties as a result of emergent behavior that often involves evolution, adaptation, or learning.

The dynamic network architectures that will be explicitly dealt with in the project are *peer-to-peer* (P2P) and *grid* computing systems [11, 3], as well as *ad-hoc networks* (AHN) [5]. P2P systems are distributed systems based on the concept of resource sharing by direct exchange between *peer* nodes, in the sense that all nodes in the system have equal role and responsibility [11]. Exchanged resources include content, as in popular P2P document sharing applications, and CPU cycles or storage capacity, as in computational and storage grid systems. P2P systems exclude any form of centralized structure, requiring control to be completely decentralized. In AHN, heterogeneous populations of mobile, wireless devices cooperate on specific tasks, exchanging information or simply interacting informally to relay information between themselves and the fixed network [12]. Communication in AHN is based on multi-hop routing among mobile nodes. Multi-hop routing offers numerous benefits: it extends the range of a base station; it allows power saving; and it allows wireless communication, without the use of base stations, between users located within a limited distance of one another.

We have chosen these domains for their practical importance in future distributed computing technologies as well as their potential for benefiting from our results. P2P/Grid systems can be seen as dynamic networking at the application level, while AHN results from dynamic networking at the system level. In both cases, the topology of the system typically changes rapidly due to nodes voluntarily joining or leaving the network, due to involuntary events such as crashes and network partitions, or due to frequently changing interconnection patterns.

The use of CAS techniques derived from nature in the context of information systems is not new. Numerous studies have abstracted principles from biological systems and applied them to network-related problems, primarily routing. However, much of the current work in this area can be characterized as *harvesting* — combing through nature, looking for a biological system or process that appears to have some interesting properties, and applying it to a technological problem by modifying and adapting it through an enlightened trial-and-error process. The result is a CAS that has been empirically obtained and that appears to solve a technological problem, but without any scientific explanation of why.

BISON proposes to take exploitation of CAS for solving technological problems beyond the harvesting phase. We will study a small number of biology-inspired CAS, applied to the technological niche of dynamic networks, with the aim of elucidating principles or regularities in their behavior. In other words, BISON seeks to develop a rigorous understanding of why a given CAS does or does not perform well for a given technological problem. A systematic study of the rules governing good performance of CAS offers a bottom-up opportunity to build more general understanding of the rules for CAS behavior. The ultimate goal of the BISON project is then the ability to synthesize a CAS that will perform well in solving a given technological task based on the accumulated understanding of its regularities when applied to different tasks. In addition to this ambitious overall objective, BISON has more concrete objectives to obtain robust, self-organizing and self-repairing solutions to important problems that arise in dynamic networks at both the system layer and the application layer. Here we outline these objectives.

Status of the Project and Participants

The BISON project will officially begin on January, 2003 and will last for three years. Given the interdisciplinary nature of the problem ahead, the BISON consortium brings together experts from a wide range of areas including core disciplines (physics, mathematics, biology) and "user" disciplines (information systems, telecommunication industry). The Department of Computer Science of the University of Bologna (Italy) brings its experience in the development of fault-tolerant distributed systems and, more recently, in the development of P2P systems Telenor (Norway's leading telecommunication company) conducts research towards nature-inspired techniques for network monitoring and surveillance, and for traffic control and network reconfiguration. The Department of Methods of Innovative Computing, University of Dresden (Germany) is an interdisciplinary institution dedicated to the modeling of biological systems and to the development of algorithms for bioinformatics and operates at the interface of computer science, mathematics and biology. The Dalle Molle Institute for Artificial Intelligence (IDSIA, Switzerland) has a long experience in the definition of combinatorial optimization algorithms based on *swarm intelligence* and *ant colony optimization*. The Santa Fe Institute (Santa Fe, USA) is considered a top-level center for studies in complexity sciences, and its participation will be uniquely valuable in achieving

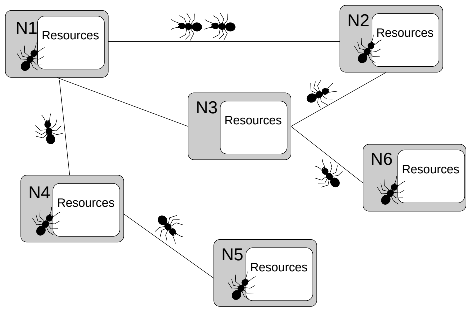

Figure 1. An Anthill System

the objectives of BISON.

Anthill

Despite the fact that the Bison project has not yet officially started, in order to pursue these ideas we have developed *Anthill*[1], a novel framework for P2P application development based on ideas such as multi-agent systems and evolutionary programming borrowed from CAS [4]. The goals of Anthill are to provide an environment that simplifies the design and deployment of P2P systems based on these paradigms, and to provide a "testbed" for studying and experimenting with CAS-based P2P systems in order to understand their properties and evaluate their performance. An Anthill system is composed of a collection of interconnected *nests*. Each nest is a peer entity that makes its storage and computational resources available to swarms of *ants* – autonomous agents that travel across the network trying to satisfy user requests. During their life, ants interact with services provided by visited nests, such as storage management and ant scheduling.

Having developed a first Anthill prototype, we are now in the process of testing the viability of our ideas regarding P2P as CAS by developing common P2P applications over it [1, 10]. For example, Messor is a load-balancing application for grid computing [10] aimed at supporting the concurrent execution of highly-parallel, time-intensive computations, in which the workload may be decomposed into a large number of independent jobs. The computational power offered by a network of Anthill nests is exploited by Messor by assigning a set of jobs comprising a computation to a dispersed set of nests. To determine how to balance the load among the computing nodes, Messor uses an algorithm inspired by the following observations. Several species of ants are known to group objects (e.g., dead corpses) in their environment into piles so as to clean up their nests. Observing this behavior, one could be mislead into thinking that the cleanup operation is being coordinated by some "leader" ant. It is possible to describe an artificial ant exhibiting this very same behavior in a simulated environment, following three simple rules: (i) wanders around randomly, until it encounters an object; (ii) if it was carrying an object, it drops the object and continues to wander randomly; (iii) if it was not carrying an object, it picks up the object and continues to wander. Despite their simplicity, a colony of these "unintelligent" ants is able to group objects into large clusters, independent of their initial distribution in the environment.

It is possible to consider a simple variant (the inverse) of the above artificial ant that drops an object that it may be carrying only after having wandered about randomly "for a while" without encountering other objects. Colonies of such ants try to disperse objects uniformly over their environment rather than clustering them into piles. As such, they could form the basis for a distributed load balancing algorithm.

Figure 2 illustrate how the load balancing process performed by Messor evolves over time. The results were obtained in a network of 100 idle nests, initially connected to form a ring (for visualization reasons). Initially, 10,000 jobs are generated in a single node. The different histograms depict the load observed in all the nests (x-axis) after 0, 5, 10, 15, 20, and 50 iterations of the algorithm. At each iteration, a set of 20 ants perform a single step, i.e. they move from one node to another, possibly moving jobs from overloaded nests to underloaded nests. As the figure illustrates, only 15-20 iterations are required to transfer jobs to all other nodes in the network, and after 50 iterations, the load is perfectly balanced. The first iterations are spent exploring the neighborhood in the network. After a few iterations, new connections are created and used to transfer jobs to remote parts of the network.

References

Additional information can be found at the project web site, `http://www.cs.unibo.it/bison`. Information about the Anthill toolkit can be found at `http://www.cs.unibo.it/anthill`.

References

[1] Ö. Babaoğlu, H. Meling, and A. Montresor. Anthill: A Framework for the Development of Agent-Based Peer-to-Peer Systems. In *Proceedings of the 22th International Conference on Distributed Computing Systems (ICDCS)*, Vienna, Austria, July 2002.

[2] E. Bonabeau, M. Dorigo, and G. Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, 1999.

[3] I. Foster and C. Kesselman, editors. *The Grid: Blueprint for a Future Computing Infrastructure,*. Morgan Kaufmann, 1999.

[4] D. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, 1989.

[5] Z. Haas and M. Pearlman. The zone routing protocol (ZRP) for ad-hoc networks (InternetDraft). Mobile Ad-hoc Network (MANET) Working Group, Aug. 1998.

[6] J. Holland. *Emergence: from Chaos to Order*. Oxford University Press, 1998.

[7] J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the Natl. Acad. Sci. USA*, 79:2554–2558, 1982.

[8] S. Kirkpatrick, C. Gelatt, and M. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[9] E. Klarreich. Inspired by Immunity. *Nature*, 415:468–470, Jan. 2002.

[10] A. Montresor, H. Meling, and O. Babaoğlu. Messor: Load-Balancing through a Swarm of Autonomous Agents. In *Proceedings of the 1st Workshop on Agent and Peer-to-Peer Systems*, Bologna, Italy, July 2002.

[11] A. Oram, editor. *Peer-to-Peer: Harnessing the Benefits of a Disruptive Technology*. O'Reilly & Associates, Mar. 2001.

[12] C. Toh. *Ad Hoc Mobile Wireless Networks: Protocols and Systems*. Prentice Hall, 2002.

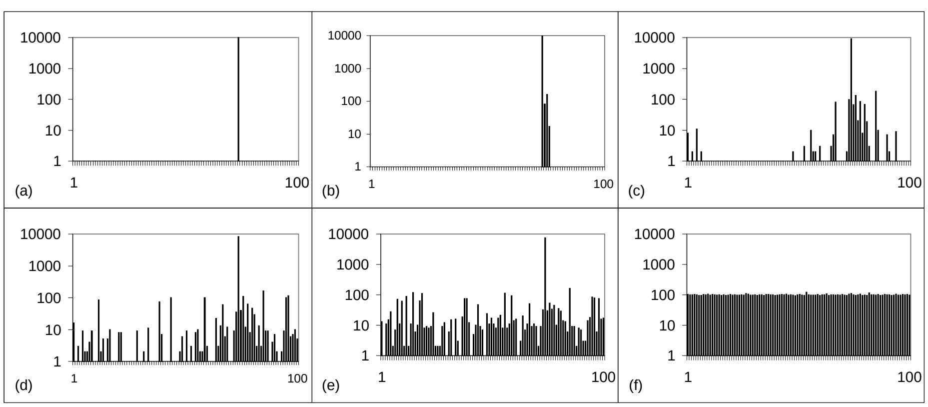

Figure 2. Load distribution after 0, 5, 10, 15, 20, 50 iterations.