Performance Evaluation of Multiagent Systems: Communication Criterion

Faten Ben Hmida, Wided Lejouad Chaari, and Moncef Tagina

University of Manouba, National School of Computer Studies, SOIE / LI3 Laboratory,
2010 Manouba, TUNISIA
{Faten.BenHmida, Wided.Chaari, Moncef.Tagina}@ensi.rnu.tn

Abstract. Many MultiAgent Systems (MAS) have been developed in various application domains such as computer networks, Internet, industrial applications, automation, process control, air traffic, robotic, simulation, etc. In spite of the rapid growth of the international interest in MAS field and the importance of the number of developed MAS, there is still a lack related to their performance evaluation. In fact, there is no measurement tool that allows evaluating the performance of a MAS or comparing two MAS. The existing works on systems performance evaluation deal principally with classic computer systems. In this paper, we try to identify the MAS' special features and properties which have an impact on their performance and we propose a measurement model to evaluate one of these properties: communication. This model is based on the graph theory. An experimentation of the proposed evaluation model is carried out and tested on a diagnosis application.

Keywords: Multiagent Systems, Communication, Performance Evaluation, Measurement, Methodology, Graph Theory.

1 Introduction

With the continuous extension of MAS application fields, there are more and more issues concerning particularly standardization of many concepts governing the development of this new technology. General issues deal with MAS terminology and incite scientists to agree on a consensus defining basic notions relative to MAS. Other issues are concerned about more precise questions such as the enumeration of the MAS's common characteristics and their estimation within the framework of evaluation prospects [1][2].

Now let's focus on the latter point. Consider the situation where we are faced to a given problem and we have several computer systems destined to resolve this problem. In this case, how could we know the most appropriate one? It is essential that we have a mean allowing us to choose the most appropriate system while we are convinced of the reasons of our choice. Yet, we should refer to a standard methodology to evaluate and analyze the performance of each system and thus make our choice.

N.T. Nguyen et al. (Eds.): KES-AMSTA 2008, LNAI 4953, pp. 773–782, 2008.

However, the problem with the MAS is that there is no mean to evaluate the efficiency of a MAS or to compare two MAS. In fact, performance evaluation was rarely considered within the framework of MAS [3]. Consequently, until now, we don't have a standard method to evaluate and quantify the common MAS's characteristics.

Our work precisely deals with this point. It is question of establishing a methodology to analyze and measure the characteristics of MAS from an organizational point of view. In this paper, we don't look on all these characteristics at the same time, but we start by establishing a mean to evaluate one characteristic which is *communication*.

This paper is organized as follow: In section 2, we briefly present some basic notions relative to system performance evaluation. In section 3 we explain our approach to evaluate MAS communication. Our evaluation system's architecture is described in detail in section 4. Section 5 shows the results obtained after the experimentation of our approach on a multiagent application. Results are followed by some interpretations and explanations. A conclusion and a look at future work are presented in section 6.

2 System Performance Evaluation

Performance evaluation is the most critical part in any system construction. It consists in determining the different aspects of a system performance and estimating them quantitatively and/or qualitatively. There is no general definition of a performance metric, it is system dependent and its definition requires understanding the system and its usage well [4]. According to Jain, it's an art of computer systems engineering [5]. There are three basic techniques through which performance evaluation can be performed [6]:

- **Analytical modeling:** consists in using abstract model based on mathematical notions to describe certain aspects of the real system. This model is analyzed numerically to evaluate the real system's performance.
- **Simulation:** consists in implementing a simplified model that reproduces the real system behavior in software.
- **Measurement:** consists in fitting the system with specific instruments that allow picking up the relevant values in order to measure the system's performance.

Most of the evaluation methodologies present in the literature take inspiration from Jain's one [5]. This methodology is composed of the following steps:

1. Define the evaluation objective and determine the system components which will be considered during the evaluation process.
2. List all the services provided by the system and the possible results following the solicitation of each service.
3. List the system criteria useful to the performance evaluation.
4. Select the system parameters having an impact on its performance.
5. Select the factors to make vary in order to observe the consequences of their variation on the system performance.

6. Choose the appropriate evaluation technique.
7. Fix the system workload (the number and the nature of requests submitted to it).
8. Carry out the experimentation.
9. Analyze and interpret the obtained results.
10. Display the conclusions to the user.

3 Evaluation Methodology

To evaluate the MAS's performance, two modifications were made on Jain's methodology in order to adapt it to this context: The first modification consists in reorganizing the different steps. In fact, we noticed that some steps are independent from the application domain. We gathered these steps together in a first phase that we call *choice phase*. The second phase that we call *implementation phase* includes the steps to realize in the presence of the application to evaluate. The second modification concerns steps 3, 4 and 5 of Jain's methodology. Here we confuse parameters, criteria and factors and we propose to identify the common *characteristics* of the MAS. Thus, the evaluation methodology we use is described below:

- First phase : Choice phase (application independent)
  - Define the objective and determine the system components.
  - List the system's characteristics.
  - Select the characteristic to evaluate.
  - Choose the evaluation technique to apply.
- Second phase : implementation phase (application dependent)
  - List all the system's services and the possible results.
  - Fix the system's workload.
  - Carry out the experimentation.
  - Analyze and interpret the obtained results.
  - Display conclusions to the user.

First of all, the evaluation objective must be the clearest possible, so that, the evaluation process deals only with the pertinent data necessary to attend this objective. We have identified many abstraction levels according to which performance data can be classified:

- The lowest level concerns data relative to the effects that the MAS has on the host computer system such as CPU utilization or memory consumption.
- The highest level concerns generic data relative to the MAS's proper characteristics which are generally used by designers to evaluate the efficiency of a design approach or to compare different design approaches.

Our objective is to evaluate the performance of MAS from an organizational point of view. This supposes to deal with the highest abstract level data. So, we aren't interested in material constraints and hardware implementations of the MAS, we are rather interested in the common characteristics of MAS. According to a study made by Boissier and al. in [2], MAS have 13 characteristics which are: Autonomy, Distribution, Decentralization, Communication, Interaction, Organization, Situation in an environment, Openness, Emergence, Adaptation, Delegation, Personalization, and

Intelligibility. The second point we should tackle is the fact of defining the system components which will be considered during the evaluation. In this work, the whole MAS is considered. However, we don't take care about the internal architectures of the several agents.

Communication is one of the most important characteristics of MAS. In fact, it is a central aspect, on the basis of the agents' interaction, and essential to realize the social attribute of the MAS [7]. Thus, in this work, we focus on the evaluation of the different aspects of communication in MAS. For this purpose, two evaluation techniques are used: analytical modeling and direct measurements on the system.

4 Communication Evaluation System

The evaluation system proposed is composed of two main modules: an observation module and a measurement module (see Fig.1 below).

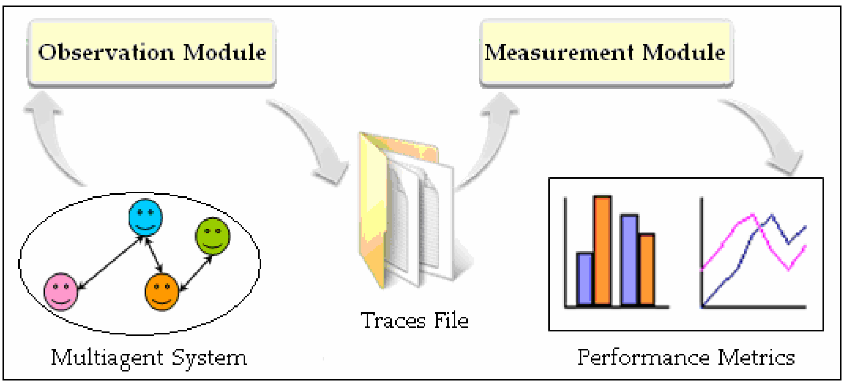


Fig. 1. Global architecture of the evaluation system

4.1 Observation Module

There are three different ways in which a MAS could be observed:

- The first consists in observing the system implicitly, i.e. spying it without interfering with it, so that there is no impact on its performance.
- The second consists in observing the system explicitly. This can be performed thanks to check points, captors or counters, etc.
- The third consists in submitting requests to the system to test its performance.

The two latter methods suppose to interfere with the system under study and consequently, the obtained results could be distorted. That's why we propose to observe our system activity using the first method i.e. implicitly, so that we don't need to analyze the effect of the observation on the system performance.

4.2 Measurement Module

Communication could be evaluated according to three main aspects which are:

- The structural aspect concerning the network communication topology.
- The syntactical aspect concerning the messages typology and complexity.
- The statistical aspect concerning the quantification of usual data such as the number of exchanged messages and their sizes.

Our study covers these three aspects. However, greater importance is accorded to the communication structural aspect. This is due to the nature of the MAS model used. In fact, our system is modeled by an oriented graph where the nodes represent the system agents and the arcs represent the communication links between those agents. Each arc is weighted by the number of messages exchanged. According to [8], an oriented graph $G = [X, U]$ is composed of:

- A set X of nodes, where $|X| = N$.
- A set U of oriented node pairs called arcs, where $|U| = M$.

There are several possible representations of a graph. Here, the adjacency matrix is used.

$$A = \left(A_{ij}\right)_{\substack{i=1..N \\ j=1..N}} \tag{1}$$

In an adjacency matrix, each line (column) corresponds to a node in the graph.

$$A_{ij} = 1 \quad \text{if and only if} \quad (i, j) \in U \quad \left(A_{ij} = 0 \text{ otherwise}\right). \tag{2}$$

In order to refine this model, each arc is weighted by the number of messages exchanged by the corresponding nodes. For this purpose, another matrix called *weight matrix* is used.

$$P_{ij} = p(u) \quad \text{if and only if} \quad u = (i, j) \in U \quad \left(P_{ij} = 0 \text{ otherwise}\right). \tag{3}$$

Structural properties of the communication graph

Once the MAS's communication graph is generated, we have to analyze its properties. Our analysis is based on graph theory which is a branch of mathematics concerned about networks encoding and properties [9]. To apprehend the graph communication structure, we start by measuring the two half-degrees of each node.

- <u>External half-degree</u> $d^{+}(n)$: it is the number of links coming from n. It reflects the participation degree of the corresponding agent to the communication act.
- <u>Internal half-degree</u> $d^{-}(n)$: it is the number of links coming into n. It reflects the solicitation degree of the corresponding agent to the communication act.

$$d^{+}(n)=\sum_{j=1}^{N} A_{nj} \quad and \quad d^{-}(n)=\sum_{i=1}^{N} A_{in} \tag{4}$$

There are other criteria useful to describe the graph's structure, namely, *indices*. Indices where developed by K. J. Kansky in 1963 in order to evaluate transport networks [9]. Some of Kansky's indices are used, which are:

- Beta index β: it is expressed by the relationship between the number of links (M) and the number of nodes (N). The higher is β the more complex is the network. So, Beta index reflects the complexity of the communication network connecting the agents.

$$\beta=\frac{M}{N} \tag{5}$$

- Gamma index γ: it is expressed by the relationship between the number of observed links and the number of possible links. Its value is between 0 and 1. A value of 1 indicates a completely connected network. In our case, Gamma index allows us to measure the communication degree in the MAS.

$$\gamma=\frac{E}{N^{2}} \text{ where } E=\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij} \tag{6}$$

- Theta index θ: it measures the average amount of traffic per node. The higher is θ the greater is the load of the network.

$$\theta=\frac{Q(G)}{N} \text{ where } Q(G)=\sum_{i=1}^{N}\sum_{j=1}^{N} P_{ij} \tag{7}$$

Also, the load $Q(n)$ of each node can be measured.

$$Q(n)=\sum_{\substack{i=1 \\ i\neq n}}^{N} P_{in}+P_{ni} \tag{8}$$

- Connectivity: A graph is said to be connected if for all its distinct pairs of nodes there is a linking chain. The study of the communication graph's connectivity enables us to have an idea about the organization of the MAS's agents. Identifying the different connected sub-graphs comes to identifying the different acquaintances in the MAS. To this end, Tarjan's algorithm [8], which allows finding the connected components in a graph, is used.
- Articulation points: In a connected graph, a node is said to be an articulation point if its suppression increases the number of connected components in the graph. The existence of such nodes in the MAS's communication network reflects some centralization in the communication. Tarjan's algorithm [8] is used to find the articulation points in the communication graph.

Syntactical properties of the communication

There are two criteria according to which the syntactical aspect of the communication is evaluated. These criteria are:

- Messages typology: The communication between agents is based on speech act theory which treats communication as action and claims that speech acts could change the state of the world just like physical actions [10]. All agent communication languages define a list of performatives corresponding to various communication acts. To study the messages typology, the performative field is extracted from each message and the number of the performatives used is counted.
- Messages complexity: Depending on the message structure, to each captured message one of these three qualifications (*simple*, *medium*, *complex*) is attributed. The qualification attributed depends on whether the content of the message is:
  - A string: in this case, the content is *simple* since it does not need to be encoded and decoded to be interpreted.
  - An ontology based clause: in this case, the content is *medium* since it is encoded according to a common ontology and needs to be decoded by the agent to be interpreted.
  - A protocol based message: in this case, the content is *complex* since a protocol specifies a number of rules and behaviors to be performed in addition to the common ontology.

Statistics

The analysis of the communication is completed by some statistics. In fact, we count the total number of exchanged messages, the messages sizes, the number of agents involved in the communication, and the percent of agents involved in the communication.

5 Experimentation

Our evaluation system has been tested on a multiagent application designed in order to detect and localize failures in an industrial system [11][12]. This application was implemented on the JADE multiagent platform. It consists of the following agents:

- The detection agents D1, D2, D3, D4 and D5, whose role is to detect the failures.
- The localization agent LOC, whose role is to localize the failure.
- The interface agent INT, whose role is to coordinate the other agents processing's and to display results to the application user.

As we mentioned before, our evaluation system is composed of two modules: the observation module and the measurement module: The observation module uses a Jade spy agent called *Sniffer* [13] to collect the several messages exchanged by the application's agents. All the messages captured by the Sniffer are saved in a traces file. The measurement module takes this file as input. An ACL parser parses it and progressively fills a table with the information contained in this file. Then this table is used to draw the communication graph and to calculate the performance metrics. Fig.2 shows the communication graph of the diagnosis application.

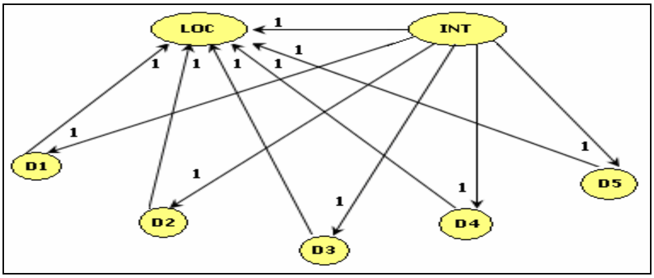


Fig. 2. The multiagent system's communication graph

According to the communication graph, we notice that the detection agents D1-5 don't communicate with each other. They rely on the localization agent LOC which is solicited by all the other agents. This is due to the fact that while detecting a failure, the detection agents send the residues' values to the localization agent. On the contrary, the agent INT sends messages and doesn't receive any one. This arises from the fact that it is an organizer agent; it transmits to the detection agents a diagnosis request and informs the localization agent to receive the residues values from the detection agents. Table 1 illustrates the several performance metrics values obtained.

Table 1. Performance metrics

Property	*Measure*	*Value*
network communication complexity	Index β	1.5714
communication degree	Index γ	0.2245
network mean traffic	Index θ	1.5714
connected components	*NBcc*	1
articulation points	*NBpa*	0
number of agents involved in communication	*NBac*	11
percent of agents involved in communication	*Pac*	100%

According to the metrics values presented in Table.1, the communication network is not complex (β index is low). γ is nearer to 0 from 1, so the MAS is characterized by a low degree of communication. In addition, the whole system is composed of only one acquaintance (only one connected component) in which all the agents participate to the communication (100% of the agents are implicated in the communication).

According to results illustrated by Fig.3, Fig.4 and Fig.5, we notice that the load isn't equitably shared and that the agents LOC and INT are involved in the communication more than the others.

Like it is shown in Fig.6, the message typology is poor; the agents use only one type of messages which is represented by the communication act INFORM. This seems to be natural according to the application needs. In fact, the communication between the several agents consists in an exchange of information; there are no sophisticated or complex conversations. According to Fig.7, the messages' sizes are almost the same. All the messages' contents have a medium complexity given that the agents send encoded matrices in accordance with the ontology defined by the application designer.

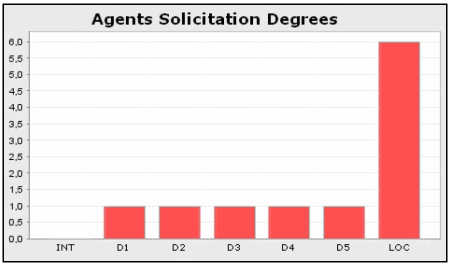


Fig. 3. The agents' solicitation degrees

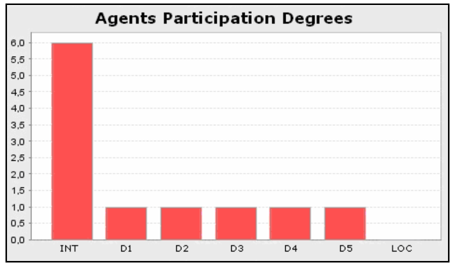


Fig. 4. The agents' participation degrees

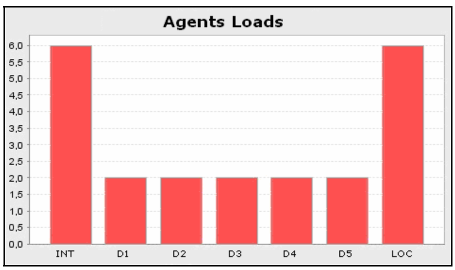
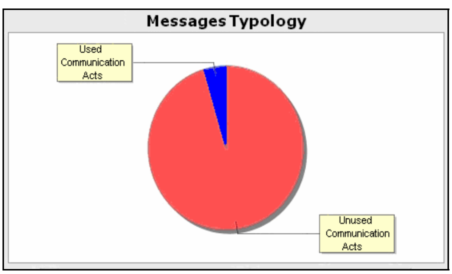


Fig. 5. The agents' loads



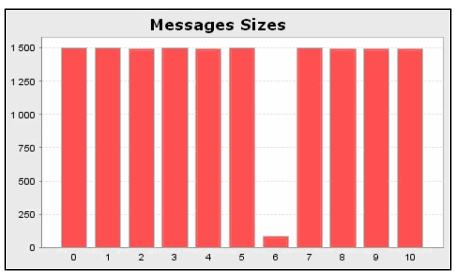


Fig. 6. Messages' sizes

Fig. 7. Messages Typology

6 Conclusion

In this paper, an experimental approach to evaluate communication in multiagent systems was proposed. The implemented evaluation tool is composed of two modules: observation module and measurement module. The observation module consists in a spy agent that captures messages exchanged by the agents and saves them in a file. This file is then exploited by the measurement module to draw the communication graph and to calculate the performance metrics. Our evaluation covered several aspects of the communication in MAS which are: the structural, the syntactical and the statistical ones. The obtained results allowed us to validate our approach. In the future, we will focus more on the observation module and we will try to find a more generic solution to the MAS's activity observation. In addition, we will be interested in some other characteristics of the MAS.

References

1. Boissier, O., Guessoum, Z., Occello, M.: Un essai de définition de critères pour l'étude comparative de plates-formes multi-agents. In: Technique et Science Informatiques, vol 21. Hermès, Paris (2002)
2. French Observatory of Advanced Technologies: Systèmes Multi-Agents. Arago 29, pp. 27–44. Tec & Doc, Paris (2004)
3. Jurasovic, K., Jezic, G., Kuzek, M.: A Performance Analysis of Multiagent Systems. International Transactions on Systems Science and Applications 1(4) (2006)
4. Le Boudec, J.Y.: Performance Evaluation of Computer and Communication Systems. EPFL (2007)
5. Jain, R.: The Art of Computer Systems Performance Analysis. John Wiley & Sons, England (1991)
6. Willig, A.: Performance Evaluation Techniques, Fundamentals and Big Picture. Telecommunication Networks Group, Technical University Berlin (2005)
7. Lejouad-Chaari, W., Mouria-Beji, F.: High-Level Communication Protocol in a Distributed Multiagent System. In: Moonis, A., Mira, J., de Pobil, A.P. (eds.) IEA/AIE 1998. LNCS, vol. 1415, pp. 606–616. Springer, Heidelberg (1998)
8. Gondran, M., Minoux, M.: Graphes et Algorithmes., Eyrolles, Paris (1995)
9. Graph Theory: Measures and Indices, http://people.hofstra.edu/geotrans/index.html
10. Wooldridge, M.: An Introduction to MultiAgent Systems. John Wiley & Sons, England (2002)
11. Saddem, R.: Conception et Implémentation d'un Système de Diagnostic à Base d'Agents. Research Master Report. National School of Computer Studies, Tunis (2006)
12. Bouabdallah, B.S., Saddem, R., Tagina, M.: A Multiagent Architecture for Fault Detection and Isolation. In: 1st International Workshop on Applications with Artificial Intelligence. Patras (2007)
13. JADE Sniffer Online Documentation, http://jade.tilab.com/doc/tools/sniffer/index.html