# Multi-agent Architecture for Heterogeneous Reasoning under Uncertainty combining MSBN and Ontologies in Distributed Network Diagnosis

Álvaro Carrera
*Dpto. de Ingeniería de Sistemas Telemáticos*
*Universidad Politécnica de Madrid, UPM*
*Madrid, Spain*
*Email: a.carrera@dit.upm.es*

Carlos A. Iglesias
*Dpto. de Ingeniería de Sistemas Telemáticos*
*Universidad Politécnica de Madrid, UPM*
*Madrid, Spain*
*Email: cif@dit.upm.es*

***Abstract*—This article proposes a MAS architecture for network diagnosis under uncertainty. Network diagnosis is divided into two inference processes: hypothesis generation and hypothesis confirmation. The first process is distributed among several agents based on a MSBN, while the second one is carried out by agents using semantic reasoning. A diagnosis ontology has been defined in order to combine both inference processes.**

**To drive the deliberation process, dynamic data about the influence of observations are taken during diagnosis process. In order to achieve quick and reliable diagnoses, this influence is used to choose the best action to perform. This approach has been evaluated in a P2P video streaming scenario. Computational and time improvements are highlight as conclusions.**

***Keywords*- agent, Bayesian, ontology, diagnosis, network**

## I. Introduction

The complexity of telecommunication networks has increased the demand for network and service management systems. Nowadays, network fault management requires high skills engineers, which are not able to cope with the increasing heterogeneity and complexity of the network. The probability of occurrence of faults in large telecommunication networks grows as they become widespread, complex and heterogeneous [1]. Thus, the role of automatic diagnosis modules is getting more attention, in order to cover the detection, isolation and recovery of faults. Another important aspect to point out is the need for dealing with uncertainty during the diagnosis task, since many corroboration tasks cannot be carried out because of different reasons, such as the cost itself of the action or that the action requires to access the subscriber equipment and could cause him any trouble.

The main focus of this paper is to present a Multi-Agent System (MAS) architecture that combines two reasoning processes: semantic reasoning and Bayesian reasoning. This approach proposes to use Bayesian inference to handle uncertainty inherent in any diagnosis process and semantic inference to discriminate which action is the best one to perform depending on the available data.

The reminder of this article is structured as follows. First, Section II proposes an agent architecture for reasoning during both phases of a diagnosis: hypothesis generation and hypothesis confirmation. These two phases can be deployed in different agents. Section III explains the case study where our approach is applied. Section IV shows the evaluation and presents the results of comparison with other approaches. Finally, Section V draws out the main conclusions about the application of this approach and, besides, a brief description of future possible improvements.

## II. Agent Architecture

This section proposes an agent architecture for diagnosis tasks which combines two reasoning processes. It consists of the following modules shown in Figure 1:

*1) Bayesian Module:* is a Bayesian reasoning inference engine that processes environment data (test results, symptoms, etc.) to infer possible root causes of the symptoms with an associated confidence about these beliefs. The outcomes of this module are hypotheses with its respective confidences and strength of influence among nodes [2].

We propose to use dynamic data from Bayesian networks. Each node of a Bayesian network has a concrete influence over its neighbours [2] in each moment of the diagnosis procedure. This influence between nodes is the strength of the dependencies between nodes that is quantified via conditional probability tables (CPT). This influence changes dynamically depending of the evidences of the network (in other words, depending of the available information about the environment).

To obtain these data, CDF [3] distance is used. This method is suitable when there are ordinal nodes, because it represents the shift of probability according to the cumulative probability functions of the two distributions.

*2) Ontology-based Reasoning Module:* has the aim of deliberating which action should be performing out of the information from the Bayesian module. Bayesian module generates an ordered list of possible actions to confirm the diagnosis hypothesis. This module filter this action based on the action preconditions. After executing the actions, the result is fedback to the Bayesian module.The reasoning process which includes rules for multimedia failure diagnosis are expressed with SWRL [4] and OWL [5].

*3) Agent Control Module:* follows an extended BDI agent architecture where beliefs are distributed across the two above mentioned inference modules. When the agent receives new symptoms, a diagnosis plan is launched, combining the previous modules.

*4) Mapping module:* translates information between *Bayesian module* and *Ontology module*. It performs the mapping process to create ontology individuals and extract information from ontology concepts to probabilistic data that can be input in the Bayesian module. To perform this task, we use PR-OWL [6] ontology that supports a way to add probabilistic information to others concepts defined using OWL.

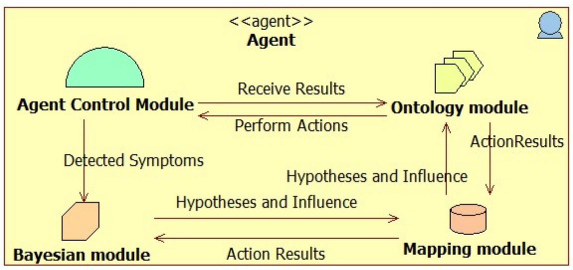


Figure 1. Agent Architecture

*Agent Types*

We have to remark that it is not necessary that all agents have all modules. Some functionalities can be distributed across several agents in order to obtain more scalability, remote access to restricted data, less computational requirements, etc.

At this point of the explanation and to clarify the multi-agent system proposed, three types of agents can be discriminated:

- *Fully Autonomous Agent* which has all modules presented before. It is able to evaluate the environment, reason (in a distributed way) under uncertainty, perform actions, etc. It can work autonomously, but it has better performance working together with other agents.
- *Semi-Autonomous Agent* which has Agent Control Module and Ontology-based Reasoning Module. It cannot deal with uncertainty, but it is able to interact with its environment. To reason with uncertainty, it has to interact with an *Fully Autonomous Agent*.
- *Dependent Agent* which has only the Agent Control Module. It is able only to perform prefixed request actions. For example, the execution of one test or one monitoring action.

The usage of multi-agent technology for diagnosis tasks in telecommunication networks brings a range of benefits. For example, agents can be deployed in remote nodes, work when they are isolated or even create other agents dynamically when its functionality is required. These features are highly recommended for systems that work in complex environments. Our proposal consists of defining a flexible agent architecture which integrates the previously identified modules. These functionalities can be distributed at design time or even run time by the agents themselves (creating agents on demand), depending on non functional requirements (time to repair) or functional requirements (distribution requirements because of actions on remote equipments).

## III. Case study

### A. Scenario

To properly frame this study, a P2P streaming scenario was chosen. In this scenario, there are a multimedia provider user and a multimedia consumer user. Many faults may occur both in connection and in services. The system is designed to provide, to an end-user or an operator, the result of the diagnosis made upon receipt of a notification of a symptom of failure. The result is expressed in percentages representing the certainty of the occurrence of a given hypothesis.

The scenario network topology is as follows:

- Multimedia Provider Home Area Network that feeds the multimedia content.
- Multimedia Consumer Home Area Network that consumes the streaming service.
- ISP intranet that belongs to the service provider.
- Access network that provides access to home users.

Sharing of multimedia contents between two home users is addressed. These contents are stored in a video server inside of the Multimedia Provider HAN (Home Area Network) and are remotely accessed from the Multimedia Consumer HAN. Multimedia contents are transmitted in real time using RTSP (Real Time Streaming Protocol) for session establishment and RTP (Real-time Transport Protocol) for content delivery.

### B. Streaming Diagnosis Case

In order to simplify the explanation of the proposed approach, only a simple case is exposed in the following paragraphs. First of all, the multi-agent system is presented, and then, an overview of the Bayesian Network as well as the design principles are discussed. And finally, the key processes of the case are highlighted.

Agents have been deployed according to geographic distribution. Thus, one *Semi-Autonomous Agent* has been deployed into the multimedia client PC. This agent has monitoring capabilities. One *Fully Autonomous Agent* and one *Dependent Agent* have been deployed into Consumer Home Gateway. These agents have diagnosis and test capabilities. Two other agents like these (one *Fully Autonomous Agent* and one *Dependent Agent*) have been placed into ISP network. And finally, one agent of each type is deployed into Multimedia Provider HAN like into Multimedia Consumer HAN. Two agents into Home Gateway one *Fully Autonomous Agent* and one *Dependent Agent* and another one into Streaming Server (one *Semi-Autonomous Agent*). Each one of these agents know which actions are able to perform by itself and publish them to allow to other agents the request of action.

Each *Fully Autonomous Agent* has its own piece of the whole MSBN, its own subnetwork. These Bayesian Networks have been modelled following the BN3M model [3] (see Figure 2). In this model, three types of variables are distinguished: *context*, *fault* and *evidence*. *Context* variables model the environment, in this case, these variables are used to model information about the network in which each agent resides. *Fault* and *Evidence* variables are used to model the possible failures through hypotheses and observations.

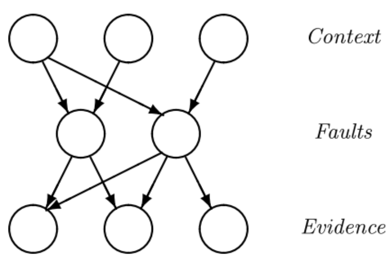


Figure 2. BN3M model

In the following, a diagnosis scenario is described based on the previously described configuration. First, a streaming session is detected by the *Semi-Autonomous Agent* that resides inside multimedia client PC. This agent performs a monitoring action to know the quality of the session. If there is a quality degradation, a symptom is generated. However, this agent has not enough information to process this symptom, and needs to cooperate with a *Fully Autonomous Agent* (in this case, the *Fully Autonomous Agent* that resides in the Multimedia Consumer Home Gateway).

This agent is able to process symptoms performing Bayesian inference in a distributed way (using MSBN approach). In other words, this agent shares information with others *Fully Autonomous Agent* that are able to reasoning with high level data. At this point, all *Fully Autonomous Agents* are working together and in parallel. Each one take its own decisions using shared and its own knowledge.

But, once a symptom has been processed, which action is the best one to perform? Depending on the state of the environment and the knowledge base of the agent, one action could change its influence in the diagnosis process. To deal with this issue, we use CDF method (see Section II). With this method, all possible actions are ordered by relevance to reach a reliable confidence in the diagnosis process. The first one whose preconditions are fulfilled is selected and executed.

Finally, when an hypothesis has a confidence higher than a threshold, the diagnosis finishes and a healing action is searched to fix the problem. But this is other issue that is not evaluated in this paper. Anyway, it is important to remark that this is a key component to close the autonomic control loop (self-recovery functionality).

## IV. Evaluation

The benefits of the proposed meta-model have been evaluated comparing this approach with previous works [7], [8]. In this paper, we compare the performance of the system using deliberation driven by "cost" or by "influence".

In previous works, test actions were classified by estimated cost. This cost combined time cost and computational cost and is estimated a priori by human experts. Then, all test actions are executed always in the same order. And sometimes, unneeded actions are executed.

The evaluation has been carried out based on a benchmark for a real diagnosis scenario of the R&D project Magneto. With data stored in data base with old diagnosis and the same Bayesian networks have been used in both cases. The volume of this data is around 500 diagnoses. We have clustered all possible diagnosis in 13 diagnosis cases to simplify comparison and shown results.

As it is shown in Figure 3, the number of performed tests has been reduced. Taking data from data base mentioned above, the average of performed test with deliberation driven by cost is 5.23 tests (with standard deviation 3.11). Using deliberation driven by influence, this number is reduced to 2.76 (with standard deviation 1.42); in other words, the number of performed tests has been reduced in 47.05%.

With deliberation driven by influence, there are two diagnosis cases that performs one test more than following the previous approach (driven by cost). The reason of this behaviour is that these are connectivity failures inside user HAN. These failures are very uncommon; for this reason, these hypotheses have, a priori, a little confidence and other hypotheses have to be confirmed or refused first.

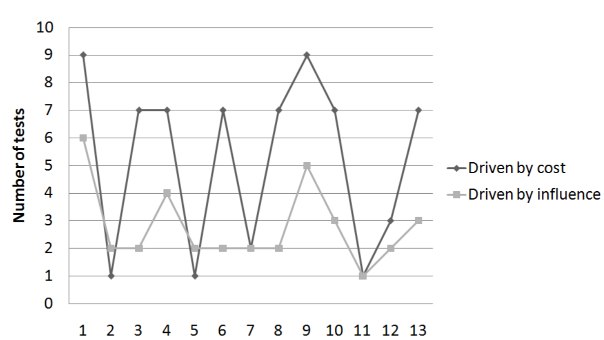


Figure 3. Comparison: previous work vs proposed approach

Table I
MTTD AND NUMBER OF TEST COMPARISON

| Diagnosis case | MTTD | | Number of tests | | Result |
|---|---|---|---|---|---|
| | Cost | Influence | Cost | Influence | |
| Case 1 | 41 | 29 | 9 | 6 | △ |
| Case 2 | 5 | 7 | 1 | 2 | ▽ |
| Case 3 | 36 | 8 | 7 | 2 | △ |
| Case 4 | 39 | 15 | 7 | 4 | △ |
| Case 5 | 4 | 8 | 1 | 2 | ▽ |
| Case 6 | 36 | 8 | 7 | 2 | △ |
| Case 7 | 9 | 9 | 2 | 2 | ∼ |
| Case 8 | 34 | 8 | 7 | 2 | △ |
| Case 9 | 40 | 24 | 9 | 5 | △ |
| Case 10 | 33 | 13 | 7 | 3 | △ |
| Case 11 | 5 | 5 | 1 | 1 | ∼ |
| Case 12 | 13 | 8 | 3 | 2 | △ |
| Case 13 | 36 | 14 | 7 | 3 | △ |

Table I shows the evaluation results in several columns. MTTD [9] (Mean Time to Diagnose) usually is the average number of minutes until the root cause of the failure is correctly diagnosed; but, in this table, we show this time rounded in seconds. Other relevant times, like MTTR [9] (Mean Time to Respond) or MTTF [9] (Mean Time to Fix), are not covered in this study.

The column named "Result" represents if deliberation driven by influence improves driven by cost one or not in a specific diagnosis case.

The average of MTTD in previous approach is 25.47 seconds (with standard deviation 15.33), in proposed approach is 12.01 seconds (with standard deviation 7.12). Time improvement is 52.87%.

## V. CONCLUSIONS AND FUTURE WORK

We have presented a MAS that uses a meta-model ontology to diagnosis with Bayesian reasoning using OWL and SWRL to choose actions to perform. We focused on the deliberation process, leaving outside other research scope issues. Our proposal of decision support improves previous approaches [7], [8] both in time and in computational cost.

As future work, we will study in depth the application of Multiply Sectioned Bayesian Networks (MSBN) [10], [11] to distribute the Bayesian inference engine that offers support to handle uncertainty and to maintain coherence and consistency in a distributed reasoning process. Applying MSBN approach, we can have a distributed inference engine that does local information processing, partial intermediate information exchange, inference global consistency and self-organization due to partial damage.

## ACKNOWLEDGEMENT

This research has been partly funded by the Spanish Ministry of Science and Innovation through the projects Ingenio Consolider2010 AT Agreement Technologies (CSD2007-0022), Cenit THOFU (CEN-20101019) and T2C2 (TIN2008-06739-C04-03/TSI) as well as the Spanish Ministry of Industry, Tourism and Trade through the project RESULTA (TSI-020301-2009-31).

The authors would also like to thank to J. García-Algarra, J. González-Ordás, P. Arozarena and R. Toribio for their support and collaboration in this research through the *Magneto R&D project*.

## REFERENCES

[1] G. Berthet and N. Fischer, "A unified theory of fault diagnosis and distributed fault management in communication networks," in *Proceedings of IEEE. AFRICON '96*. IEEE, pp. 776–781.

[2] U. B. Kjaerulff and A. L. Madsen, *Bayesian Networks and Influence Diagrams*, ser. Information Science and Statistics. Springer New York, 2008.

[3] P. Kraaijeveld, M. Druzdzel, A. Onisko, and H. Wasyluk, "Genierate: An interactive generator of diagnostic bayesian network models," in *Proc. 16th Int. Workshop Principles Diagnosis*. Citeseer, 2005, pp. 175–180.

[4] M. O'Connor, H. Knublauch, S. Tu, B. Grosof, M. Dean, W. Grosso, and M. Musen, "Supporting rule system interoperability on the semantic web with SWRL," *The Semantic Web - ISWC 2005*, pp. 974–986, 2005.

[5] D. Q. Z. Xiao Hang Wang, *Ontology based context modeling and reasoning using OWL*. IEEE, Mar. 2004.

[6] P. C. G. Costa and K. B. Laskey, "PR-OWL: A framework for probabilistic ontologies," in *Proceeding of the 2006 conference on Formal Ontology in Information Systems Proceedings of the Fourth International Conference FOIS 2006*. IOS Press, 2006, pp. 237–249. [Online]. Available: http://portal.acm.org/citation.cfm?id=1566079.1566107

[7] A. Sedano-Frade, J. González-Ordás, P. Arozarena-Llopis, S. García-Gómez, and A. Carrera-Barroso, *Distributed Bayesian Diagnosis for Telecommunication Networks*, ser. Advances in Intelligent and Soft Computing, Y. Demazeau, F. Dignum, J. M. Corchado, and J. B. Pérez, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, vol. 70.

[8] F. J. García-Algarra, P. Arozarena-Llopis, S. García-Gómez, and A. Carrera-Barroso, "A lightweight approach to distributed network diagnosis under uncertainty," in *INCOS '09: Proceedings of the 2009 International Conference on Intelligent Networking and Collaborative Systems*. Washington, DC, USA: IEEE Computer Society, 2009, pp. 74–80.

[9] J. FitzGerald and A. Dennis, *Business Data Communications and Networking*. John Wiley and Sons, 2008.

[10] Y. Xiang, D. Poole, and M. P. Beddoes, "Multiply sectioned bayesian networks and junction forests for large knowledge-based systems," *Computational Intelligence*, vol. 9, no. 2, pp. 171–220, May 1993.

[11] Y. Xiang, "Belief updating in multiply sectioned Bayesian networks without repeated local propagations," *International Journal of Approximate Reasoning*, vol. 23, no. 1, pp. 1–21, Jan. 2000.