KOWLAN: A Multi Agent System for Bayesian Diagnosis in Telecommunication Networks

Sergio García-Gómez, Javier González-Ordás, F. Javier García-Algarra, Raquel Toribio-Sardón, Andrés Sedano-Frade
Telefónica I+D, Spain
{sergg,javiord,algarra,raquelt,andress}@tid.es

Ferrán Buisán-García
Telefónica España, S.A., Spain
fernando.buisangarcia@telefonica.es

Abstract—The classical approach for network diagnosis systems is based on centralized solutions with deterministic algorithms. However, the increased complexity of new networks and services, including customers premises, demands new strategies to face two key challenges: scalability of the system and reliability against incomplete or inaccurate information. This paper describes KOWLAN, a Multiagent system in which agents carry out Bayesian inference to drive the diagnosis process and to obtain diagnosis results. The system has been quickly developed and successfully deployed, and it’s contributing to a reduction of the attendance time.

Keywords—Network Diagnosis; Bayesian Networks; Multiagent Systems; Ontologies

I. INTRODUCTION

The management of telecommunication services and networks is a critical issue for a telco operator. From the conception and planning of a service to the quality assurance tasks, there are a wide range of, some times unknown, activities [1]. The management systems tenets and architectures have evolved together with the network changes and the telecom market.

For years, the main approach was to build better and more complex systems to support the tasks of technicians. These systems are usually called Operating Support Systems (OSS). However, in order to reduce operating expenses, telecommunications operators try to reduce the need of human intervention in the management processes. With this aim, in recent years the emerging field of autonomic communications proposes the definition of networks and systems that manage themselves [2].

A common management process is the diagnosis of telecommunications networks and services. Whenever a failure is detected, it is reported to a trouble ticketing system, which serves as a task queue for the technical departments in charge of solving the breakdown. Two of the main drawbacks of current diagnosis systems are scalability and management of incomplete or inaccurate information [3].

Traditionally, network diagnosis systems have followed a centralized approach and are based on remote access to information and test capabilities to come up with a conclusion. Networks have evolved towards larger and more complex scenarios, but this has not been an important issue until now. Customer premises are becoming more complex, both in networked companies and in the digital home, and telecom operators need to manage the customers’ domain. Integrating the own network diagnosis with such a huge number of networks will require a new strategy [4]. KOWLAN project proposes to deploy an overlay network of intelligent agents, which share information both in the telco networks and in the customers’ networks.

Another important drawback of current diagnosis systems is the inability to tackle incomplete, vague or inaccurate information [5]. There is a wide range of information problems that prevents classical deterministic diagnosis systems to reach valid diagnosis results: reliability problems of network inventories, observation errors, lack of testing resources, etc. With the use of bayesian inference as a reasoning mechanism inside intelligent agents, KOWLAN provides the best diagnosis results based on the available information.

The final objective of the system is to automate human tasks as much as possible. The diagnosis results provided, together with their certainty, enables KOWLAN to implement simple dispatching rules even in this early stages.

The second section of the article describes the scenario in which KOWLAN concepts have been applied, and the main challenges are explained. In the third section, there is an overview of the architecture and how the system has been developed, paying special attention to the agents ontology and Bayesian Networks. Finally, the main conclusions about the usage of the system and the challenges are remarked, together with the description of future evolution paths.

II. SCENARIOS

The KOWLAN approach is being used in different scenarios. The purpose of this article is to show how it is being developed and deployed to diagnose troubles and automate responses in the Telefonica’s MacroLAN service.

MacroLAN is Telefonica’s solution to build Virtual Private Networks (VPN) that connect multiple enterprise sites over Ethernet based accesses. It supports service speeds from 2 Mbit/s to 1 Gbit/s. By using standard L2 and L3
VPNs, MacroLAN enables geographically distant customer sites to communicate as if they belong to the same LAN, in terms of speed, reliability and transparency.

The MacroLAN service is built on top of a diverse set of Telefonica networks. In the local loop from the customer site to the Central Office, it can use either fiber or copper. Ethernet converters or Synchronous Digital Hierarchy (SDH) circuits allow to extend the distance of this access segment. Copper based accesses may also rely on Telefonica’s leased line network (named Ibermic). MacroLAN traffic is then first aggregated into province-wide Metropolitan Ethernet networks (MANs) and then into an IP backbone network that provides national coverage.

There are two main challenges that diagnosis for the MacroLAN service shares with most telecommunication network diagnosis processes: incomplete information and scalability.

A. The incomplete information challenge

There are many reasons why information is uncertain for a diagnosis system [3]. Topology and inventory data is scattered among multiple systems and it is not always complete neither consistent. The complexity of the network environment makes very difficult or even impossible to observe certain aspects of the domain, and moreover, the relationships between domain events are not always deterministic, particularly between observable and nonobservable ones. Besides, it is often impractical to model and analyze explicitly all the dependencies.

When there are failures, the behaviour of the systems is uncertain since they have been conceived to work deterministically. The observations that can be obtained from the domain, when it is feasible to get them, are sometimes inaccurate or vague. Other times, it is not possible to get the observations because there are not enough observation resources or they are not accessible and the state of the network or the service must be guessed from partial information.

In such scenario, a deterministic system such as those based on rules or on diagnosis processes is not able to get accurate results in every situation. In KOWLAN, different Bayesian Networks have been defined in order to properly diagnose each of the above mentioned scenarios, which have been defined based on the knowledge of the expert technicians.

B. The scalability challenge

The MacroLAN service is provided on the operator’s infrastructure and resources. The classical diagnosis processes usually entail getting observations from these resources. With this objective, in KOWLAN different specialized agents have been created in order to remotely gather evidences from the network and the systems.

However, one of the biggest challenges is to take into account the customers’ infrastructure in the diagnosis process. In KOWLAN, it is possible to deploy agents in the remote network, giving visibility over the devices and their configuration, and getting evidences from that environment that can be shared with the rest of the agents. This highly distributed approach enables the system to solve local problems without the participation of common resources and provides a solution to the scalability handicap of more centralized designs.

III. Architecture and development

A. System Architecture

The system architecture consists of three main blocks: a multiagent platform, a web user interface and the database. Regarding the core of the system, several types of agents have been created, but two of them are of special relevance: on the one hand, the diagnosis agents are specialized for each topology scenario and perform the bayesian inference with the evidences gathered from other agents; on the other, different observation agents obtain evidences by connecting to network equipments, databases (MIBs), and executing tests (e.g., a ping between routers).

There are other types of auxiliary agents created with different purposes:

- To receive requests from the web interface.
- To get problems from the trouble ticketing systems.
- To access other data sources and get inventory data.
- To make diagnosis results persistent in a database.

1) Bayesian Networks: The basis for the behaviour of KOWLAN agents are the Bayesian Networks. As it has been mentioned, there is one specific Bayesian Network defined for every topology scenario. Together with the causal graph, KOWLAN Bayesian Networks include additional information that help the diagnosis agents to drive their process.
• The hypothesis of the problems that the system is able to diagnose, including the certainty threshold used to stop the process.
• The tests and evidences that the different agents can obtain and their cost in terms of execution time and required resources. They also include the information that the observation agents need to get the evidence, as the network inventory details.

These Bayesian Networks are defined with the Genie editor and translated into the KOWLAN ontology so that they can be used and shared by the agents. An specific Knowledge Agent is used to redistribute the bayesian knowledge among agents when it changes.

2) Ontology: Sharing uncertain information and managing probabilistic knowledge requires to design and define a special ontology as lingua franca among intelligent agents, from which it is possible to rebuild and feed with observations the different Bayesian Networks. This ontology consists of two blocks of knowledge: the Bayesian Network structure and the current diagnosis operation information.

The bayesian structure knowledge (figure 2) includes diagnosis hypothesis (problems), observers (available evidences), dependencies and conditional probabilities and the additional information for the diagnosis process. The diagnosis operation knowledge (figure 3), includes the current observations details, and actual beliefs about every hypothesis.

3) Diagnosis procedure: Once a diagnosis operation starts, several steps are given:
• Gather all the inventory information about the circuit: topology scenario, equipment data and configuration.
• Select the appropriate diagnosis agent based on the scenario.
• Select the best available test to perform, taking into account its cost and if there is enough data to execute it. Then, request it to the observation agent, get the results, perform the bayesian inference and check if any hypothesis threshold has been reached. If not, select another test.
• When there is enough certainty or there are not more tests to perform, carry out the required action, if any, and store the results into the database.

B. System Development

KOWLAN development is based on Open Source tools. The most important one, and the basis for the multiagent architecture is WADE/JADE. JADE is a complete and sound multiagent Java platform based on FIPA standards led by Telecom Italia. Following JADE principles [6], KOWLAN agents are implemented using behaviours and FIPA-ACL messages. WADE environment also provides an important mechanism to manage the deployment of the system agents.

Although there are more sophisticated ways to represent probabilistic knowledge with ontologies as [7], KOWLAN ontology described in previous section has been implemented using the formalism envisaged by JADE. It is not based on a heavy logical background, but it copes with the system requirements in this stage, including an efficient serialization mechanism into FIPA-SL language. Moreover, the first trials implemented with OWL-DL/RDF as exchange language among agents showed up serious efficiency problems for the purposes of this application.

1http://genie.sis.pitt.edu/
2http://jade.tilab.com/
3http://jade.tilab.com/wade


The practical implementation was carried out with Ontology Bean Generator\(^4\), a Protégé plugin for JADE. The ontology was populated with the data read from the Bayesian Networks (Genie files), the agents interpret this information to drive the diagnosis process as explained and use the ontology to share information about the diagnosis operations [8].

The SAMIAM\(^5\) bayesian inference library was chosen because it is a Java native implementation. When bayesian inference procedure is triggered, a Bayesian Network is constructed based on the working ontology of the agent, and the available evidences for the given diagnosis operation. The beliefs obtained from the probabilistic reasoning process are then incorporated into the ontology, so that they can be shared with other agents or stored into the database.

Although the system is intended to be triggered by the Trouble Ticketing System, there is a web user interface from which the technicians can launch diagnosis operations and look up all the results. The system provides a detailed view of the diagnosis, including the customer main information, an ordered list of beliefs and their certainty, all the observations performed and the inventory data.

IV. CONCLUSIONS AND NEXT STEPS

KOWLAN system for MacroLAN scenario is a continuous development based on Scrum principles [9]. The version explained in this paper corresponds to the first six months of development. Around once a month, a new version is released and shown to the users. In such demos, the user’s feedback is collected and next steps defined.

Currently, more than 100 incidences are automatically treated per day. Around 30% of them are automatically dispatched towards other technician units, contributing to a significative reduction of the response time. In the rest of the cases, the automatic diagnosis results are very valuable to prioritize the incidences, since it is more efficient to first diagnose the a-priori easier ones.

The first challenge for KOWLAN system is its capacity to work with incomplete or inaccurate data. In this scenario, it has been observed that many times there are problems getting information and observations from different sources. The probabilistic approach provides an answer in any case, although the higher the number of evidences, the higher the certainty the system gets. There is a number of automatic diagnosis in which the results, not being conclusive, provides valid clues to the technicians to diagnose the problem.

Regarding scalability, it has been seen that the multiagent paradigm provides a good solution to easily deploy decoupled pieces of software in an environment where the overall architecture is not known. The agent’s can proactivity

\(^4\)http://protege.cim3.net/cgi-bin/wiki.pl?OntologyBeanGenerator
\(^5\)http://reasoning.cs.ucla.edu/samiam/

detect and solve local problems, avoiding the participation of more centralized agents of the system, which is the only chance of successfully manage the service of such amount of customers.

There are several future lines for the evolution of KOWLAN. First, it is foreseen to advance towards an autonomic communications approach, by creating new agents which would be able to automatically fix the more simple diagnosed failures.

KOWLAN already provides a tool that allows the users to give feedback to the system about the validity of a diagnosis result. With this information, learning algorithms will be applied to improve the Bayesian Networks structure and conditional probabilities.

KOWLAN approach is also been applied to different scenarios, specially related to the digital home and the mobile network. The challenge will be to deploy JADE agents in the users’ premises, as the Home Gateway, and to manage the diversity and complexity of such environments.

Finally, there are other challenges as the dynamic generation of Bayesian Networks based on an ontological description of the services and network resources, instead of using different Bayesian Networks for each application scenario.

REFERENCES


