VSAT network troubleshooting – capturing knowledge and using it

by Lars Moltsen[1], Raquel Barco[2], Pedro Lazaro[2] and Sasikanth Munagala[1]
1) Wirtek A/S, Nibevej 54, DK-9200 Aalborg SV, Denmark
2) University of Malaga, E.T.S.I.Telecomunicación, 29071 Málaga, Spain

Today the main part of the troubleshooting process is manual, but ATSIG, a research project headed by Danish software provider Wirtek in cooperation with Siemens Austria, has now been defined under the framework of a cooperation agreement with the European Space Agency, to develop a concept for automation of troubleshooting processes.

Running a Satcom network is quite a challenge: A huge amount of extremely remotely located pieces of equipment have to play together to ensure satisfaction of the customer who is expecting a highly reliable end to end data connection. However, equipment can break down or can be incorrectly set up, a thunderstorm can cause poor signal quality to/from a ground station, or external radio sources can interfere with the carriers. These examples illustrate the complexity that the Satcom operator faces.

The Troubleshooting Process

The Satcom troubleshooting process is similar to troubleshooting in many other domains. It can be decomposed into the following three parts:

1. **Fault Detection**: Identifying that some part of the network has a problem (yet not knowing what the problem is).
2. **Diagnosis Creation**: Investigation of the problem to identify the root cause (what is the problem?).
3. **Solution Deployment**: Once the root cause of problems has been identified, a solution has to be deployed onto the system.

The level of automation in the Satcom operator troubleshooting process is typically limited to the *Fault Detection* part, in which alarm systems and automated scripts normally help monitoring staff to detect when the network behavior is abnormal. The rest of the process relies on human experience and knowledge to couple alarms, measurements, and other pieces of information to identify possible root causes and to carry out solution deployment. The level of automation is shown in *Figure 1*.

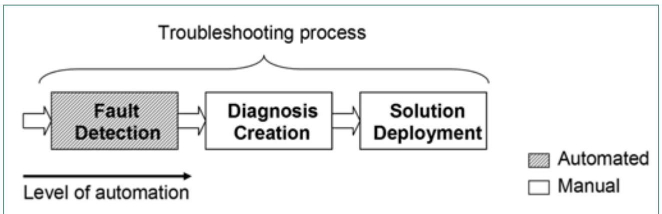

Figure 1: Current Satcom operator troubleshooting process

The reason why the *Diagnosis Creation* task has not been successfully automated is due to system complexity: It is not possible to use typical diagnostic solutions like rule-based systems or decision trees, since the level of uncertainty is simply too high for these "deterministic" techniques to work. Too often, an input/symptom is not available, or due to randomness it is "on" when it would normally be "off". In such cases, a decision tree will either get stuck (missing value) or it will follow a wrong branch and come up with a wrong conclusion (random behaviour).

The ATSIG project aims to solve automation of the *Diagnosis Creation* task by utilizing a technique from the artificial intelligence (AI) domain known as Bayesian networks. This technique is very robust towards missing input data, and has a built-in method for handling randomness. It has shown to be quite useful for other complex diagnostic purposes including both medical diagnostics as well as troubleshooting of computers, printers, terrestrial radio networks, and huge electrical locomotives. The resulting solution in ATSIG will increase the level of automation in the troubleshooting process to also cover *Diagnosis Creation* as shown in *Figure 2*.

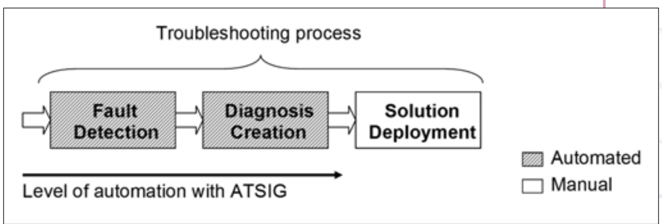

Figure 2: Troubleshooting process using the ATSIG auto-diagnosis concept

Solution Requirements

In order to automate the *Diagnosis Creation* task, it is necessary to analyze current processes and define a number of requirements that the solution should meet. The following list has been identified:

1. *The system shall be able to diagnose a selected subset of root causes with a similar (or better) quality compared to a human troubleshooter given the same input.*
   This is the basic requirement of an automated diagnosis system. Given some input in the form of a set of *symptom* readings, the system shall be able to provide an output which is at least as good as what the human troubleshooter could achieve with the same input. However, it will make sense to limit the amount of pos-

sible faults/root causes that the system can handle to a subset covering somewhere in the order of 90-99% of all troubleshooting cases. The reason is that the remaining root causes are rare and thus the time spent to model these in the system may not be worth the effort as long as the system is able to output that the problem is something else than the covered root causes.

2. *The system shall replicate human knowledge transfer.* Imagine a human troubleshooter who is working to solve a troubleshooting case and has already analyzed the problem based on the immediately available symptom readings. Then, for some reason, he needs to go home and hands over the case to a colleague. The handover will typically consist of a statement like: "My analysis shows that the most likely problem is X because of …, but it could also be Y due to …". The automated diagnosis system should be able to replicate this type of communication.
3. *Human experts shall be enabled to store knowledge about diagnosis creation.* Since the key knowledge about how to identify the root causes today is in the brains of the Satcom operator troubleshooting engineers, the system should provide an easy means for capturing such knowledge in a knowledge base/model.
4. *New knowledge shall be stored and used to improve accuracy.* When the automated diagnosis system has been taken into use, a lot of new cases will become available where both the symptom readings and eventually the actual root cause will be known. Such data is new knowledge/experience, which should be used to improve the knowledge base.

The proposed solution

In April 2007, the first prototype of the ATSIG system was available, and some trials and demo sessions were conducted to validate functionality against requirements.

Figure 3 shows the main diagnosis output window of the ATSIG system, where a diagnosis has been computed for a carrier on the Aalborg site where a "Low EIRP" alarm has indicated that some problem is present. In order to diagnose the problem, the ATSIG system has retrieved three "symptom readings" shown in the "Evidence" panel. From these, a ranking of four specific root causes plus "Other Problem" and "No Fault" is made from the computation of their *probability*.

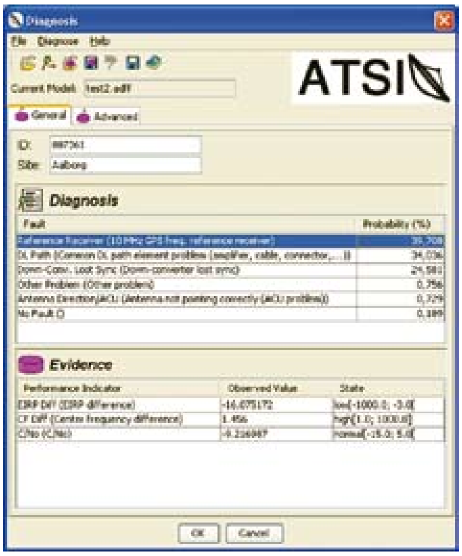

Figure 3: The ATSIG diagnosis window

The example clearly demonstrates how the ATSIG project has met Requirement 2 of "replicating human knowledge transfer": Where the human language is rich and where you can stress that the problem "is VERY likely to be X because of …", ATSIG replicates this by quantifying belief with probabilities and in addition monitoring the inputs (symptom readings). This also means that in some cases (like the presented example) more work is required by a human engineer to determine exactly which one, out of two-three root causes with high probability, is the true fault. This can be avoided by specifying many inputs with a high degree of information on specific root causes. A good set of inputs will automatically increase the average certainty of the diagnoses, such that the amount of manual interaction is kept at a minimum. The number of symptoms in this example (three) is much too low to provide value in real-life cases.

The knowledge base behind this example is defined using a special *Model Maintenance Tool*, where a troubleshooting expert has specified:

1. A set of *root causes* and a set of *symptoms/inputs*.
2. For each symptom:
   a. SQL scripts for automatically retrieving the actual symptom readings/observations of the specific case
   b. A mapping of symptom values to states (e.g. "low", "normal", "high").
3. In order to compute the probabilities for the diagnosis:
   a. The prior probability of the root causes (looking at all solved cases, how frequent each root cause is)
   b. The conditional probability of each symptom given each related root cause

By enabling the user to specify this "domain model" we meet Requirement 3 in the previous section. In the ATSIG tool, the information is converted to a *Bayesian network*, which is a graphical representation of causal and probabilistic relations, and which allows for accurate computation of posterior probabilities of the root causes (what is the probability of X given that we have seen Y and Z?).

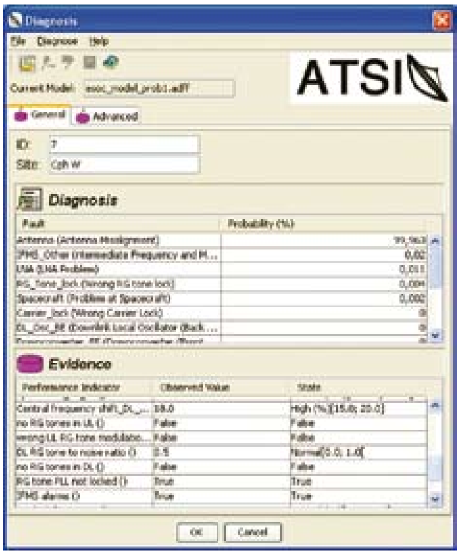

Figure 4: Another example model and data set with 22 inputs and 15 different root causes

The ATSIG system scales well although in general probabilistic reasoning has exponential complexity (poor scaling properties). The reason is that by using certain constraints on the model structure, the complexity is kept linear. *Figure 4* shows a more realistic model than the one in *Figure 3*. This model is built for satellite control diagnostics, and it has 15 possible root causes and 22 inputs.

Figure 5 shows the flow of data in the ATSIG system. The "Model" (the knowledge base) specifies to the ATSIG reasoning engine how the actual symptom readings (from alarms, events, and measurements) are coupled to root causes. The ATSIG reasoning engine produces diagnoses to the end users, and once the end user closes a case (by marking what the actual problem was), the model is updated according to this new piece of knowledge using an "adaptation" step using standard statistical calculations. This means that new knowledge is continuously captured and added to the model, which will improve over time, and this is also the direct solution to Requirement 4 in the previous section.

The ATSIG project is still in "Phase 1", where the system has been designed and the first available prototype has been demonstrated to a number of operators. In order to verify that Requirement 1 was met (diagnosis performance at least as good as a human expert), a large-scale trialing campaign is planned to be conducted in "Phase 2", running until spring 2008.

Bayesian Networks

This section presents the fundamental technology of the ATSIG project, namely Bayesian networks. Bayesian networks have been used in a range of diagnostic applications ranging from medical decision support systems to printer and PC troubleshooting (Hewlett-Packard and Microsoft) to locomotive troubleshooting (General Motors). Wirtek has also recently used this technology in TheCure, a system for troubleshooting the radio access part of GSM and UMTS networks.

A Bayesian network models the *random variables* of a domain. *Figure 6* shows a small Bayesian network, where three disease variables are present (Measles, Mumps, and Rubella). In addition, two symptom variables are also present (Fever and Spots). For simplicity, all variables could be assumed to have two states, "yes" and "no". The model also contains *causal links*. A causal link means that the state of a variable has a causal impact on the state of another variable – e.g. Measles would have a causal impact on Fever and Spots whereas Mumps only impacts Fever.

In addition to the graphical structure, the Bayesian network has one (conditional) probability table per variable. The rule is that it should be the conditional probability of the variable given each of its *parents* (incoming links). Thus, in *Figure 6* there are 5 probability tables, and e.g. the table of Spots will be: *P(Spots | Measles, Ruballa)* (the conditional probability of Spots given Measles and Rubella). For variables without incoming links, the probability table is simply the prior probability of each state of the variable – e.g. *P(Measles) = [yes: 0.01; no: 0.99]*.

Once the structure and conditional probabilities are specified, the Bayesian network is ready to be used by using a combination of *Bayes' rule* and properties of conditional independencies. Basically, any observation of a set of variables can be fed into the Bayesian network and updated posterior probabilities can be computed for all other variables. *Figure 7* shows an example where both Fever an Spots have been observed to be "yes". The effect is that the probabilities of Measles and Rubella become high (the reason why Rubella scores highest is that Fever was set to have a slightly higher correlation with Rubella, compared to Measles).

Benefits of ATSIG to the Satcom Operator

The current ATSIG prototype solution has now been demonstrated to a number of Satcom operators, both to validate the proposed solution and to reveal additional requirements of the final solution. The initial feedback has been very positive.

Bo Hjorth Jensen, the head of technology in Emperion, a provider of broadband network solutions via satellite mainly active in Europe, Middle-East, and Africa, says: "A system that can both capture knowledge and automate our processes will bring a lot of value to my organization. We will be less sensitive to the right people being at work, and we will be able to satisfy customers even better than today."

Based on the operator feedback, the following list of main benefits have been identified:

1. *Saving time and resources*
2. *Decreased time to solve problems*
3. *Better customer satisfaction*
4. *Less sensitivity to human knowledge*
5. *Knowledge is captured and stored*

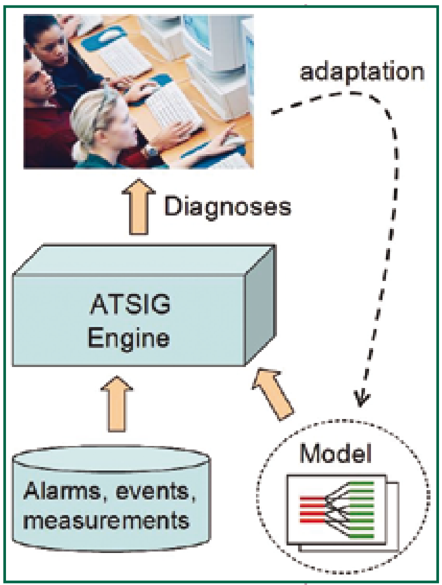

Figure 5: The ATSIG system data flow

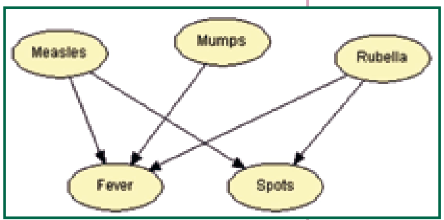

Figure 6: A Bayesian network for diagnosing child diseases

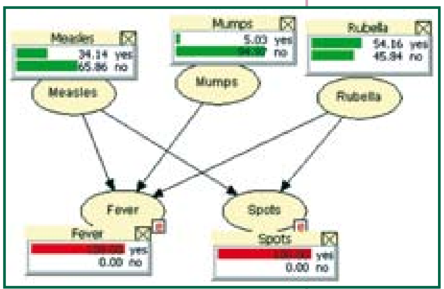

Figure 7: The diagnostic example with Fever and Spots observed to be "yes"

Summary

Although the ATSIG project is still not completed, it has already, in the prototyping phase, provided demonstration results that indicates a great potential of using probabilistic reasoning technology/Bayesian networks for automation of the troubleshooting process. And not only will the toolset help speed up and standardize the troubleshooting process in the operator organization, it will also secure knowledge within the organization and not leave the operator suffering if key personnel go on vacation or decide to leave.

In Phase 2 of the project (autumn 2007 – spring 2008), the ATSIG toolset needs to be thoroughly validated in cooperation with operators, and the usability and integration properties must be further developed in order to provide an "off-the-shelf" troubleshooting product for Satcom operators.

The project progress can be followed at www.atsig-project.org, and further information can be acquired by contacting Lars Moltsen from **Wirtek**.

Q

Peer reviewed

Lars Moltsen (+45) 25 21 46 35
or Lars.Moltsen@wirtek.com