

Using Autonomic Computing and Click Stream Analysis for Problem Identification in Continuous Production

Wallace A. Pinheiro^{1,3}, Jonice Oliveira¹, Jano M. de Souza^{1,2},
Geraldo Xexéo^{1,2}, and Marcelo Perazolo⁴

¹ Computer Science Department, Graduate School of Engineering
Federal University of Rio de Janeiro (UFRJ)
PO Box 68.511 - ZIP code: 21945-970 – Rio de Janeiro, RJ – Brazil
{awallace, jonice, jano, xexeo}@cos.ufrj.br

² Computer Science Department, Institute of Mathematics
Federal University of Rio de Janeiro – Brazil

³ Computer Systems Engineering Section - SE/8
Military Institute of Engineering (IME)

Pr. General Tiburcio, 80 - ZIP code: 22290-270 - Rio de Janeiro, RJ – Brazil

⁴ IBM Corporation – RTP, NC, USA
mperazol@us.ibm.com

Abstract. Problem identification is an area of research from Autonomic Computing. The problem identification can bring several benefits in a production line, especially in continuous production of the Information Technology Industry. This work presents an approach based on Symptom Ontologies to facilitate problem identification and solution prediction, and use of clickstream analysis for recommendation of material, when the solution is not appropriated.

1 Introduction

One of the main steps in an engineering process is the production, which is responsible for the construction of products that will reach the consumer. If a design is to be manufactured by continuous production, the factory or production line will operate continuously, without interruptions, and the product is passed from machine to machine.

If a machine stops working due to a problem, it can affect directly the productivity. All time used to fix or exchange this machine is spent time, probably incurring in some loss.

A new paradigm from IT can help to effectively detect and prevent problems associated with applications and equipments. It is Autonomic Computing, which provides concepts such as Common Events and Symptoms to facilitate automatic processing, and consequently, identification, prediction and treatment of problems and incidents. Among those, the detection and description of effective symptoms assumes a prominent role, since any action taken to correct a failure should be based on them.

Some problems are easily inferred by human experience and symptoms can then be easily described, but in other cases the huge complexity of IT solutions makes impracticable for domain experts to know beforehand what kinds of failures, and consequently symptoms, they can expect from their environment. To solve this problem we can employ system verification methodologies and possible solutions' reasoning, based on past cases (like the Case-Based Reasoning paradigm).

This kind of solution, which is based on monitoring and reasoning, needs a common and well-defined knowledge representation of the domain. For it, we use ontologies, which are defined as "a specification of a conceptualization" [1]. That is, an ontology is a formal description of the concepts and relationships that represent a domain. Therefore, for identification, prediction and treatment of problems and incidents in IT elements, one requires a Symptoms Ontology, which consists of elements that describe what symptoms and solutions exists, and what are their relationships.

Sometimes, the reasoning engine can not propose or indicate a solution. In that case a human agent (as a technician) has to solve the problem. If the problem is unusual, the best approach is to recommend some informative material (as pages of forums, web pages with some practices, technical reports and so on) of related issues. In our approach this recommendation is made using click stream analysis of a web server.

Click stream analysis involves to collect and analyze clicks made by users of a site. This analysis can report aggregate data about which pages users visit and in what order. It is the result of the succession of mouse clicks that each user makes (that is, the click stream). Because a large volume of data can be gathered through click stream analysis, many applications use data warehouses to keep and organize these data [2]. Normally, these click streams are stored in Web Server logs. There are a lot of commercial applications that focus on Web log analysis, but most of them are quite limited when the focus is the recommendation or solution analysis.

This work proposes the use of Symptom Ontologies to facilitate problem identification and solution prediction, and use click stream analysis for recommendation of material, when a solution is not satisfactory. For this purpose, some principles will be explained in section 2 and our proposal will be described in section 3. The section 4 outlines the conclusion and future works and the section 5 is dedicated to acknowledgements.

2 Symptoms and Events

An event is an occurrence related to the functioning of a computing node, e.g., a peer in a grid. A symptom, which comes from an event, is the identification that a possible problem occurred. In the Symptom Analysis approach, it may also be identified from state data, metrics or log records. For instance, when a user starts a program (event) and the computer that runs this program stops working (symptom) because of the program, we have examples of an event and a symptom caused by this event.

Events, in our contexts, are defined as actions or occurrences which a system can respond; examples include clicks, key presses and mouse movements. A symptom is a