An Autonomic Computing Approach to Problem Diagnosis in Muti-Agent Systems

Sand Luz Corrêa
Outline

• Motivation
• Monitoring Systems
• Bayesian Network
• Proposed Work
• Expected Results
• Schedule
• Bibliography
Motivation

• Service Level Objectives are kept through monitoring activities

• Monitoring: observation of the system components and their relationship

• Problem: fast growth of system size and complexity
  – Brings about a combination problem

• Solution:
  – To develop systems able to
    • self-monitoring
    • reasoning about their own state
    • Self-diagnosis
    • Self-repair
System Monitoring

- Generic Monitoring Architecture for Autonomic Systems

- User interface
- Activation Layer
- Analysis Layer
- Monitoring middleware
- Application

- Policies
- Performance
- Failure
- Dependencies
- Monitor
- Database
System Monitoring

• Monitoring Middleware
  – Physical Structure
    • Hardware, communication network
  – Operational Structure
    • Middleware ou operational system
  – Application
    • System specification, System interaction
System Monitoring

- Approaches for the Analysis Layer
  - Use Knowledge Discovery in Database technique (KDD)
  - KDD is grouped into 3 categories:
    - Symbolic: incorporates a priori model of system structure and behavior as a set of event-condition-action rules
    - Artificial Intelligence: neural networks
    - Statistical models: statistical classifiers assume that attribute values are distributed through probabilistic model.
Bayesian Network

- **Meaning**
  - A data structure to represent dependencies among variables that gives a precise specification of any full joint probability distribution

- **Used to model circumstances where:**
  - There is causal relationship among data
  - There is uncertain about the domain
  - Classify probabilistic relationship among variables from a given domain through the conditional probability stated by Bayes Rule

\[
P(A|B) = \frac{P(B|A)P(A)}{P(B)}
\]

\[
P(H|E_1,..,E_n) = \frac{P(E_1,..,E_n|H)P(H)}{P(E_1,..,E_n)}
\]
Bayesian Network

• Representation
  
  – A bayesian network is a directed acyclic graph in which
    
    • A set of random variables makes up the nodes of the network
    
    • A set of directed links connects pair of nodes. If there is a link from node X to node Y, X is said to be parent of Y
    
    • Each node Xi has a conditional probability distribution that follows:
      
      – P(Xi | Parents(Xi)) that quantifies the effect of the parents on the node
Bayesian Network

- Example

Burglary

Earthquake

Alarm

JohnCalls

MaryCalls

P(B) = 0.001

P(E) = 0.002

B  E  P(A)
T  T  0.95
T  F  0.94
F  T  0.29
F  F  0.001

A  P(J)
T  0.90
F  0.05

A  P(M)
T  0.70
F  0.01
Bayesian Network

• Inference in Bayesian Network
  – Prediction:
    • Compute the probability distribution of some nodes given the distribution of their parents
    • Inference from causes to effects
  – Diagnosis:
    • Compute the probability distribution of some nodes given the distribution of their children
    • Inference from effects to causes
  – Anomaly Detection
    • Anomalies can be detected by the likelihood value: the measure of how well the observations fit the model
    • The more the likelihood is low, the more the observation is anomalous
The Classification Pattern Problem

• Diagnosis as a classification pattern problem
• What do we want to classify?
  – If a system is in compliance (1) or violation (0) according to a SLO specification
    – \( S = \{0, 1\} \), state variable
• What will be monitored?
  – A set of variables whose values affect directly the system state w.r.t. a SLO specification
    – \( M = \{m_1, m_2, \ldots, m_n\} \), metrics variable
• We want to induce a classifier, that is, a function that can map any possible value of \( M \) to a value of \( S \).

\[ F(M) = S \quad F = P(S|M) \]
The Classification Pattern Problem

• TAN(Tree Augmented Bayesian Network)
  – Naive Bayes: puts a structure upon the network topology, making the process more efficient
  – TAN: Restricts a bayesian network to a Markov tree:
    • The state variable S is the root of the tree and the parent of any other node
    • Each metric $M_i$ has at most one parent $M_j$, other than S
Restrictions and Advantages

• Restrictions
  – The metrics must be well chosen to capture system states relating to the behavior of interest
  – The analysis must observe a statistically significant sample of instances, once the learning is based on training set.

• Advantages
  – Efficiency
  – Easy to interpret
  – Modifiability through the induction of new models
  – Correct results for studies holding few variables
Proposed Work

- **Purpose:**
  - Implement a multi-agent monitoring architecture for diagnosing timing failures, using TAN as a tool for data analysis
  - Validate the architecture with a chat case study

- **Applications**
  - Real time
  - Performance constraint

- **Variables to be monitored**
  - CPU usage
  - Memory usage
  - Calls to methods
  - Network usage (packets sent and received)
Expected Results

- Evaluate the applicability of statistical models as a building block to construct self-healing systems
- Extend the analysis layer with more expressive models, like models that can deal with changes in time.
Schedule

- Study of bayesian network foundations
- Study of tools to implement bayesian network
- Definition of the variables to be monitored (Study of timing faults diagnosis)
- Implementation of the monitoring middleware layer
- Study about data training
- Training the network with the case study application
- Result evaluation

• Adaptative Replication of Large-Scale Multi-Agent Systems – Towards a Fault-Tolerant Multi-Agent Platform. Guessoum, Z.; Faci, N.; Briot, J.P. ICSE'02 4th International Workshop on Software Engineering for Large-Scale Multi-Agent System. 2002

• Applying Feedback Control in Adaptative Replication Mechanisms in Fault Tolerant Multi-Agent Organization. Bora, S.; Dikenelli, O. SELMAS'06

• Diagnosing a Team of Agents: Scalling-Up. M. Kalech, G.A. Kaminka. AAMAS'05. 2005


• On Fault Tolerance in Law-Governed Multi-Agent Systems. Gati, M.; Lucena C.; Briot, J. P. SELMAS'06

• Exception Diagnosis in Open Multi-Agent Systems. Shah, N; Chao, K.M; James A. Intelligent Agent Technology, IEEE/WIC/ACM International Conference. 2005


Perguntas ?
Obrigada !