Distributed Case-based Reasoning for Fault Management

Ha Manh Tran and Jürgen Schönwälder
Computer Science, Jacobs University Bremen, Germany

Computer Networks and Distributed Systems

JACOBS UNIVERSITY

Research Goal and Questions

The goal of the research project is to assist operators in finding solutions for faults occurring in the large-scale, diverse communication systems of high importance. Research concentrates on three questions: *1. How to exploit various fault knowledge sources in a distributed environment for similar faults? 2. How to represent faults to better retrieve relevant information from knowledge sources? 3. How to reason on the retrieved solutions to induce new solutions for new faults?*

Case-based Reasoning

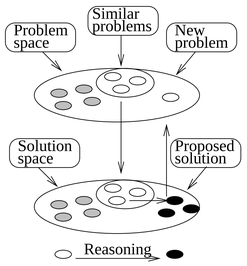

Fig.1 Case-based reasoning

Case-based reasoning (CBR) has been proposed in the early 1990s to assist operators in the resolution of faults by providing mechanisms to correlate an observed fault with previously solved similar cases (faults). Existing CBR systems for fault management usually operate only on a local case database and can not easily share and exploit knowledge about faults and their resolution present at other sites.

Distributed Case-based Reasoning

The proposed distributed CBR system exploits various fault knowledge sources in a distributed environment to retrieve similar cases, and then reasons on the retrieved solutions to induce new solutions adapted to the circumstances of a given fault.

Peer-to-Peer Architecture

The system takes advantage of P2P technologies to achieve some degree of self-organization, scalability in architecture, and flexibility in search in a decentralized and federated environment.

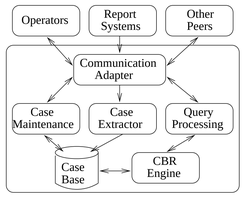

Fig.2 The architecture of distrib. CBR peers

The architecture shown in Fig.2 defines several functional blocks related to operating CBR engines, processing cases and queries, or communicating with operators, report systems and peers. A peer thus needs sufficient bandwidth and processing power to carry out all the tasks. Hence, we employ a backbone network of super peers, see Fig.3, and propose a feedback-based search mechanism, see Fig.4, for the system.

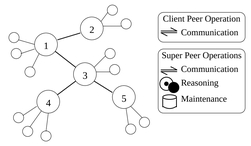

Fig.3 Super peer network overlay

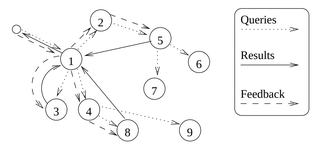

Fig.4 Feedback-based search mechanism

Fault Retrieval

The system exploits semantics-based search mechanisms associated with fault representation methods to improve fault retrieval in CBR engines.

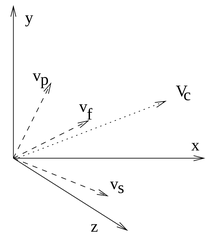

Fig.5 The multi-vector representation

We propose a multi-vector representation method to represent faults:

$$V_c = \{v_f, v_p, v_s\}$$

A field-value vector: $v_f = <f_1{:}v_1, \ldots, f_n{:}v_n>$ classifies a fault into the smaller groups of network faults. Another field-value vector $v_p = <p_1{:}v_1, \ldots, p_m{:}v_m>$ specify symptoms and typical parameters depicted in domain-specific terminology. A semantic vector v_s exhumes the properties of a case hidden in natural language. Several functions are used to evaluate similarity between fault cases:

- $cos(q, c) = \sum_{i=1}^{m} q_i c_i$ — Cosine function for semantic vectors $\|q\| = \|c\| = 1$
- $sim(q, c) = \sum_{i=1}^{n} w_i sim_{\sigma(i)}(q_i, c_i)$ — Global similarity function for field-value vectors q and c

Collaborative Reasoning

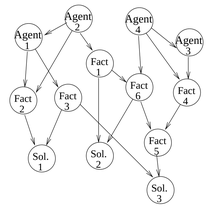

Fig.6 Collaborative reasoning capability

CBR engines only contain limited fault databases and partial knowledge about faults. The system provides a collaborative mechanism among CBR engines to augment their individual reasoning capabilities. We envision to support interactions among CBR engines to provision "ensemble" solutions.

Evaluation Methods

We use several evaluation methods for the various components of the system. The performance of the search mechanism is important in the P2P overlay because query processing in CBR engines may require significant computational resources.

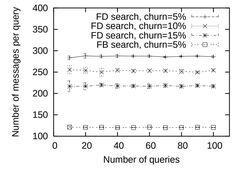

Fig.7 Reduced number of messages per query

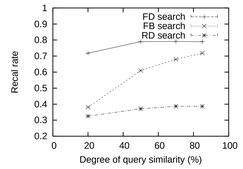

Figure 8 Recall rate of the various sets of queries

The performance of the fault retrieval closely depends on the multi-vector representation method, similarity functions and datasets. The preliminary evaluation results relies on the CISI and MED bibliographic datasets.

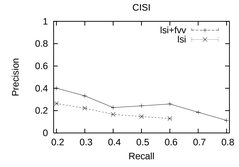

Fig.9 Precision by various recall rates for CISI

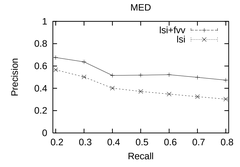

Fig.10 Precision by various recall rates for MED

Evaluation of the collaborative reasoning of CBR engines is work in progress. Besides, building a repository of realistic cases is necessary for the real evaluation of the system.

†Supported in part by the EC IST-EMANICS Network of Excellence (#26854).