Winter 12-9-2003

Discovery and Support of Problem-Solving Knowledge in e-Business

Chih-Kun Ke

Duen-Ren Liu

Discovery and Support of Problem-Solving Knowledge in e-Business

Chih-Kun Ke Duen-Ren Liu

Institute of Information Management
National Chiao Tung University
Hsinchu, Taiwan

clay.iim89g@nctu.edu.tw dliu@iim.nctu.edu.tw

Abstract

Employees generally need to solve various situations of problems occurred during task-executions in e-business. Such problem-solving behaviors contain rich information helpful to provide effective knowledge support. This work proposes a mining framework to discover problem-solving knowledge by analyzing problem-solving behaviors recorded in an intranet portal log file. The discovered knowledge can provide guidance and support to assist employees handle various situations of problems.

1. Introduction

The intranet became a platform for enterprises to deliver inner services to employees. Companies adopt Internet-based electronic services (e-services) conducting business process and sharing of process information. Database and knowledge-based system are set up as corporate knowledge repositories to manage corporate knowledge. Intranet portal provides the e-service to facilitate employees getting desire process knowledge in corporate knowledge repositories.

Providing task-relevant information to support effective task-executions in business processes has recently been investigated [1, 2, 4, 11, 16]. However, discovering task-relevant knowledge, in particular problem-solving knowledge has not been addressed. Employees generally need to solve various situations of problems occurred during task-executions in e-business. This work develops a mining framework to discover employee problem-solving knowledge from intranet access log file. The discovered problem-solving knowledge can assist employees search for solutions to provide effective task executions.

2. Related work

Kaula [8] outlined the necessary strategies for problem-solving. Wooldridge and Jennings [12, 13, 14] proposed a preliminary problem-solving model of team action. Heh [7] proposed a graphic-based evaluation model of problem-solving. Nair and Ramnarayan [10] investigated the relationship between cognitive needs of individuals and their effectiveness in solving complex problems. Gregor [5] investigated the role of explanations in knowledge-based systems to solve cooperative problems. Zhuge et al. [17] presented a cognitive-based problem-solving framework consisting of a problem-solving space, a category-based representation, and a set of problem-solving control strategies. Zhuge [16] designed an approach integrating workflow and agent concept to cooperate problem-solving process.

Synonymy and polysemy issues have created limitations in information retrieval (IR). A high level, uniform knowledge framework is needed to better enable effective solution searching. Zhuge [15] proposed an active document framework (ADF) to make knowledge documents self-representable, self-explainable and self-executable. Moreover, Liu et al. [9] proposed a semantic topic method to assist user in self-learning.

Various techniques are necessary to support the discovery of hidden knowledge in problem-solving process. Data mining is a useful tool for discovering useful information from a large amount of data. The association rule technique is a marvelous method for discovering relationships between knowledge items. Well-known mining algorithms are *Apriori* and *Apriori-like* methods [3, 6].

3. Framework overview

A mining framework to support problem-solving process is illustrated in Figure 1. This framework is designed for e-business workflow environments. When employees encounter problems, they rely on problem-solving knowledge, e.g. user guides and experiential reports. Problem-solving knowledge is stored in corporate knowledge repositories. Companies provide Intranet portal to facilitate the tracking of employees' problem-solving behaviors, e.g. relevant documents accessed during the problem solving process. The employees' access to corporate knowledge repositories is an important resource for mining problem-solving knowledge. The intranet portal log file records the relevant documents accessed by employees. Problem-solving knowledge can be discovered from the access log file. The proposed framework contains three phases, including data preprocessing, knowledge discovery and knowledge support. In the data-preprocessing phase, a formalization engine is employed to unify the representations of various kinds of knowledge document forms. A knowledge-node format is defined to uniform the knowledge representations. Moreover, an inference engine is used to identify uncertain access behavior of employees. On the basis of uniform knowledge representations, the knowledge-discovery phase employs a mining engine to discover the problem-solving knowledge. The discovered

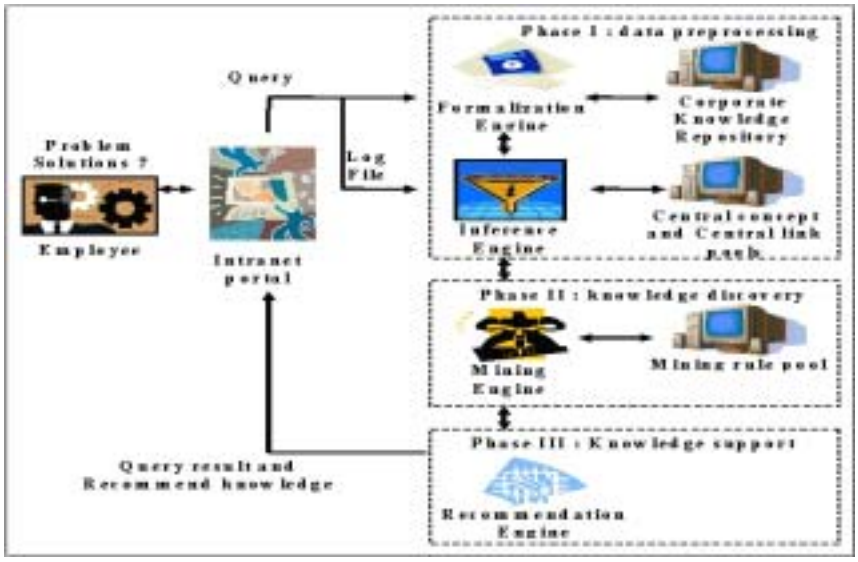

Figure 1: A mining framework to support problem-solving process

knowledge will be stored in knowledge support pools. Finally, the knowledge-support phase utilizes a recommendation engine to recommend relevant problem-solving documents to employees according to employee profiles. The recommendations assist employees solve further problems and provide effective support for task executions in e-business.

4. Proposed framework

4.1 Formalization Engine

The presentation formats of documents are varied. These can include letters, faxes, order forms, notifications, receipts, memos, etc. Although the knowledge documents are digitalized, it is hardly to search information among various types, we need a uniform format to formalize the knowledge document presentation. Knowledge document with uniform format will be the best knowledge support to solve problem occurred in e-business. In data preprocessing phase, formalization engine uses node format to provide such function.

The main function of formalization engine is to construct knowledge document profile. Semantic-based method is used to define related knowledge representation format for e-business. In order to make it easy to modularize corporate knowledge documents, we define the business knowledge document as level-wise node type:

Definition 1: *Knowledge Node* (*KN*): Each corporate knowledge document is considered as a knowledge node. *KN* is the first level.

Definition 2: *Abstract Conceptual Node* (*ACN*): Abstract conception means that expert uses a knowledge representation to abstract the concepts behind *KN*. Each knowledge document may contain one or many abstract conceptions. Each abstract conception is presented as an abstract conceptual node. *ACN* is the second level.

Definition 3: *Semantic link* (*SL*): After analyzing *ACNs'* contents, if a semantic relation is existed, a semantic link will be constructed to connect two *ACNs.*

4.2 Inference Engine

The inference engine transforms the employees' sequential access behaviors recorded in the log file into *KN* representations, which corresponds to a *KN* sequence. Notably, the information of specific problem in the *KN* level may be sparse. To gain more information for mining process, we pay attention to the *ACN* level. From *ACN* level, a *KN* sequence corresponds to an *ACN* network. In addition to *ACNs*, *SLs*, semantic relations exist in *ACN* network. We define these relations as inference links.

Definition 4*: Inference link* (*IL*): In sequential access behavior, *KN*s form a banding status. If no semantic relation exists between *ACN*s, an inference link may be constructed to connect them. Inference link is inferred as a candidate of human cognitive behavior.

For each problem, *ACNs* network is considered as a *Conceptual Knowledge Access Network* (*CKAN*). Because of various situations, specific problem may own one or many *CKANs*. By analyzing each *CKAN*, we get weight and frequency of *ACNs*, *SLs*, and *ILs*. From knowledge document profile, we get weight of *ACNs. ACNs'* information used to construct each *CKAN's* concept profile. Links' information is used to construct each *CKAN's* link profile. Proposed framework uses two pool mechanisms to store profile information. *Central concept pool* (*CCP*) is used to store all *CKAN's* concept profiles of specific problem. Higher weight or frequent *ACN* is considered as central concept. The filtering criterion is based on threshold setting. *Central link pool* (*CLP*) is used to store all *CKAN's* link profiles of specific problem. Higher weight or frequent link is considered as central link. The filtering criterion is also based on threshold setting.

Inference engine carries out inferring process during *CKAN's* concept and link profile construction. Working steps of process are illustrated as the following:

(1) Read employee access behavior in intranet portal's log file.
(2) Identify each knowledge document in access sequential patterns and search for its profile in knowledge repository.
(3) Identify *KNs* and *ACNs* to construct *CKAN* according to knowledge document profiles.
(4) Construct connecting links and identify link types of each pair of *ACNs*.
(5) Record weight and frequency information of each *ACN*, *SL*, and *IL* in *CKAN's* concept profile and link profile.

4.3 Mining engine

Data preprocessing phase provides information support for problem-solving process in e-business. In knowledge discovery phase, a mining engine is employed to discover the problem-solving knowledge. A specific problem may have more than one *CKAN* concept profile. We consider each *CKAN* concept profile as a transaction of specific problem. Concepts in *CKAN* concept profiles are considered as item in transaction. Mining engine uses mini-conf and mini-support as the thresholds of executing *Apriori* algorithm. A knowledge support pool (*KSP*) is used to store association rules of specific e-business problem. Discovered rules are considered as a specific problem's solving knowledge. The steps of mining process are illustrated as follows:

(1) Select all *CKANs'* concept profiles of specific problem in *CCP*.
(2) Consider each specific problem's *CKAN* concept profile as a transaction and central conceptual nodes as the knowledge items of the transaction.
(3) Use *Apriori* algorithm to discover association rule.
(4) Mining results stored in *KSP* are specific problem's support knowledge.

4.4 Recommendation Engine

In e-business, knowledge support plays a key role to provide effective problem-solving process. In the knowledge-support phase, a recommendation engine is designed to suggest key steps and documents for problem-solving process. A *Guiding Problem-Solving Process* (*GPSP*) is supported to assist the problem-solving process. Association among central concept nodes is used to predict empirical behavior in problem-solving process. *SL* and *IL* in specific *CKAN's* link profile will be used to connect relevant central concept nodes to construct *GPSP*. *GPSP* shows employee the empirical problem-solving steps and key knowledge documents. Two modules are designed to providet recommendations.

Identification module

- **Problem identification function**

(1) Accumulate and gather statistics of nodes occurring frequency in all *CKAN*s' concept profiles of specific problem in e-business.
(2) Use threshold methods to filter *ACNs*. Nodes with weight or frequency higher than some thresholds are considered as central concepts of specific problem solution. Central concepts will be constructed in specific problem description stored in *CCP*.

- **Problem feature matching function**

(1) Real time to catch employee access behavior in intranet portal.
(2) Accumulate and gather statistics of *ACNs* occurring frequency.
(3) Identify central concepts of current problem and use clustering method to search for matching problem descriptions in *CCP*.
(4) According to match process, it classifies current problem into specific problem or produce new problem identification.

Selection module

(1) Get the central concept nodes and mining profile of specific problem description stored in *CCP*.
(2) Get the mining profile of specific problem stored in *KSP*.
(3) Use mining profile of specific problem to identify the relations between central concept nodes.
(4) According to central concept nodes and their relations, select higher weight and frequent *SLs* and *ILs* in specific *CKAN's* link profile to construct *GPSP* for knowledge recommendation.

5. Conclusions

Problem-solving knowledge plays a key role in supporting effective e-business services. In problem-solving process, employee access behavior contains rich knowledge. This study enforces practical knowledge discovery in problem-solving process. Proposed mining framework discovers hidden knowledge of employee access behavior. It divides knowledge discovery process into three phases. In data-preprocessing phase, a well-defined format of knowledge representation is defined to support mining and information search. In knowledge-discovery phase, mining mechanism discovers context-aware problem-solving knowledge of e-business. Knowledge-support phase provides employees guidance and documents relevant to problem-solving. The guiding process shows employee the experiential problem-solving steps and important knowledge documents.

References

[1] Andreas Abecker, Ansgar Bernardi, Knut Hinkelmann, Otto Kuhn, Michael Sintek, “Context-Aware, Proactive Delivery of Task-Specific Information: The KnowMore Project”, Information Systems Frontiers, 2000, 2 (3/4), 253-276.

[2] A. Abercker, A. Bernardi, H. Maus, M. Sintek, C. Wenzel, “Information supply for business processes: coupling workflow with document analysis and information retrieval”, Knowledge-Based Systems, 2000, 13, 271-284.

[3] R. Agrawal, R. Srikant, “Fast algorithm for mining association rules”, Proceedings of twentieth International Conference on Very Large Data Bases, 1994, 489-499.

[4] Kurt D. Fenstermacher, “Process-Aware Knowledge Retrieval”, Proceedings of the 35th Annual Hawaii International Conference on System Sciences, 2002, 209.

[5] Shirley Gregor, “Explanations from knowledge-based systems and cooperative problem-solving: an empirical study”, Human-Computer Studies, 2001, 54, 81-105.

[6] Jiawei Han, Micheline Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann, 2001.

[7] Jis-Sheng Heh, “Evaluation Model of Problem-solving”, Mathematical and Computer Modeling, 1999, 30, 197-211.

[8] Rajeev Kaula, “Problem-solving strategies for open information systems”, Knowledge-Based Systems, 1995, 8(5), 235-248.

[9] Bing Liu, Chee Wee Chih, Hwee Tou Ng, “Mining Topic-Specific Concepts and Definitions on the Web”, Proceedings of the twelfth International Conference on World Wide Web, 2003, 251-260.

[10] K. Unnikrishnan Nair, S. Ramnarayan, “Individual differences in Need for Cognition and Complex Problem-solving”, Journal of Research in Personality, 2000, 34, 305-328.

[11] S. Staab, H.-P. Schnurr, “Smart task support through proactive access to organizational memory”, Knowledge-Based Systems, 2000, 13 (5), 251-260.

[12] Michael J. Wooldridge, Nicholas R. Jennings, “Formalizing the cooperative problem-solving process”, 13th International Workshop on Distributed Artificial Intelligence, 1994, 403-417.

[13] Michael J. Wooldridge, Nicholas R. Jennings, “Towards a Theory of cooperative problem-solving”, Proc. Modeling Autonomous Agents in a Multi-Agent World, 1994, 15-26.

[14] Michael J. Wooldridge, Nicholas R. Jennings, “The cooperative problem-solving process”, Journal of Logic & computation, 1999, 9 (4), 563-592.

[15] Hai Zhuge, “Active Document Framework ADF: Concept and Method”, Proceedings of the Fifth Asia Pacific Web Conference, 2003, 341-346.

[16] Hai Zhuge, “Workflow- and agent-based cognitive flow management for distributed team cooperation”, Information & Management, 2003, 40, 419-429.

[17] Hai Zhuge, Jian Ma, Xiaoqing Shi, “Abstraction and analogy in cognitive space: A software process model”, Information and Software Technology, 1997, 39 463-648.