

Abstraksi

Workflow Management Systems (WfMS) provide a fundamental technological infrastructure to define and manage business processes efficiently. WfMS logs contain valuable data that can be used to discover and extract knowledge about the execution of workflows and processes. One piece of important and useful information that can be discovered is related to the prediction of the path that will be followed during the execution of a workflow.

We call this type of discovery, path mining. In this paper, we present and describe how path mining can be achieved using different data mining techniques including the Multimethod approach.

Keywords: web processes; workflows; path mining; Multimethod approach quality of service; business process management systems; data mining.

Reference to this paper should be made as follows: Cardoso, J. and Leniè, M. (2006) 'Web process and workflow path mining using the Multimethod approach', *Int. J. Business Intelligence and Data Mining*, Vol. 1, No. 3, pp.304–328.

Biographical notes: Jorge Cardoso joined the University of Madeira (Portugal) in March 2003. He previously gave lectures at University of Georgia (USA) and at the Instituto Politécnico de Leiria (Portugal). He received his PhD in Computer Science from the University of Georgia in 2002. While at the University of Georgia, he was part of the LSDIS Lab. where he did extensive research on workflow management systems. In 1999, he worked at the Boeing Company. He was the coorganiser and cochair of the First and Second International Workshop on Semantic and Dynamic Web Processes.

He has published several refereed papers in the areas of workflow management systems, semantic web, and related fields.

Mitja Leniè's work research interests are concentrated on intelligent systems, programming languages, complex systems, chaos theory, and software quality and metrics. His research results were implemented in SQUFOL project financed by the EU. He has written many technical and research papers published in recognised international journals and major conferences and coauthored various textbooks.

I. Pendahuluan

E-Commerce dan e-Service telah berkembang sangat pesat saat ini. Pada media web menjadi pasangan antara e-commerce dan e-services, ini adalah memungkinkan sebuah jaringan ekonomi baru (Sheath et al. 1999). Ruang lingkup dari aktifitas tersebut melalui tahap proses dari intra-enterprise workflows, predefined inter – enterprise dan business to business process, sampai kepada dinamika pemrosesan web diantara pengelola organisasi. Business Process Management Systems (BPMS) menyediakan sebuah infrastruktur teknologi yang mendasar untuk menjadikan dan mengelola proses bisnis secara efisien. Sebuah BPMS merepresentasikan secara tunggal, kesatuan model, integrasi dan eksekusi lingkungan sehingga dapat menerapkan dan mengimplementasikan dari sebuah proses bisnis (Smith and Finger 2003). System tersebut dapat mengelola proses web dan mengatur cara kerja web service dan aktivitas terbaik.

Konsep dari path mining dapat digunakan secara efektif bahkan banyak aplikasi bisnis yang menggunakannya untuk memperkirakan Quality of Service (QoS) dari sebuah web proses dan workflow (Cardoso et al, 2004). Path mining lebih akurat dalam prediksi algoritma sehingga membangun, membuat dan menghitung QoS untuk proses secara otomatis berbasis pada QoS aktifitas atom.

Tulisan ini menjelaskan sangat sistematis dalam proses path mining, gambaran konsep yang mendasar tentang kepraktisan path mining. Kita dapat membedakan teknik-teknik data mining, yang didalamnya ada pendekatan multimethod (Lenic and Kokol, 2002), J48 (Weka, 2004), Naïve Bayes, Sequential Minimal Optimisation (Platt, 1999), and MultiBoost (Webb, 2000) .

II. Business processes, workflows, and WfMS

Sebuah proses bisnis merupakan sekumpulan dari aktifitas yang menggambarkan berbagai metode alternatif pekerjaan yang diperlukan suatu objek bisnis. Proses bisnis terdapat di semua departemen organisasi seperti halnya, SDM, Hukum, Keuangan , Teknik dan Operasional, Penjualan dan Pemasaran dan Departemen Pelayanan.

Contoh dari sebuah proses bisnis termasuk diantaranya Pemeriksaan Aplikasi Pinjaman, Laporan Perjalanan, Proses Klaim Asuransi.

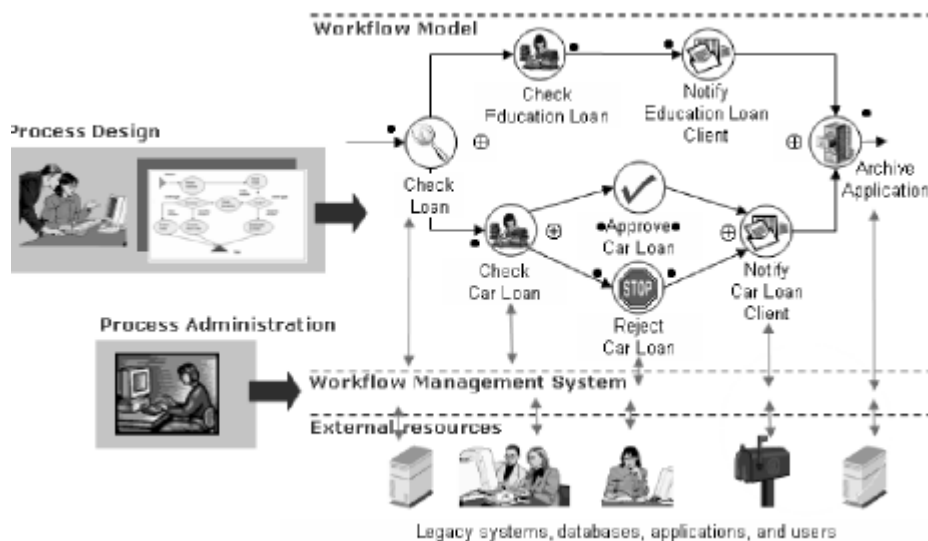
WfMS menyediakan suatu lingkungan yang dapat dijalankan untuk mengendalikan, memonitor, dan mendevlop kode yang diperlukan untuk mengotomasi bisnis proses.

Workflow merupakan proses bisnis yang otomatis dan dapat memberikan keuntungan kepada organisasi diantaranya.

1. Real Time business performance monitoring
2. Remote proses management
3. Enhanced performance and reduce bottlenecks
4. Cuts operations costs
5. Increases customer satisfaction

Dibawah ini menggambarkan komponen utama dari WfMS.

A workflow and its integration with external resources



III. Proses Path Mining dan Pendekatan Multimethod

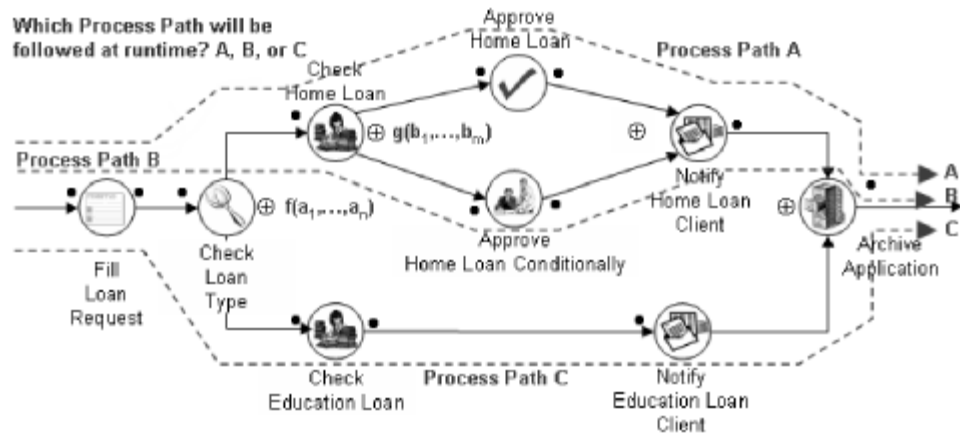
3.1. Pengenalan terhadap proses path mining

WfMS mengidentifikasi 4 jenis relasi diantara aktifitas sebagai berikut :

- a. Sequential maksudnya beberapa aktivitas dijalankan secara berurutan
- b. Paralel maksudnya dua atau lebih aktivitas dijalankan secara bersamaan
- c. Conditional, maksudnya satu percabangan yang dipilih dan diambil dari berbagai alternative percabangan workflow
- d. Iteration, maksudnya workflow dari pengulangan aktifitas yang meliputi pengulangan eksekusi satu satu lebih aktifitas workflow.

Proses path mining merupakan prosedur untuk memperkirakan urutan dari aktifitas, didalam suatu proses, yang dapat dijalankan secara langsung menggunakan mesin workflow. Urutas aktifitas tersebut diatas jika direlasikan secara langsung terhadap variable workflow. Contoh, seperti yang terlihat pada gambar 2.

Figure 2 Process path prediction



3.2. Pendekatan Multimethod

Untuk memperoleh path mining dapat digunakan 2 teknik utama :

- Analisis Stokastik
- Data Mining

Kedua teknik tersebut diatas menggunakan data yang tersimpan dalam BPMS log.

Analisa stokastik meliputi penentuan kemungkinan dari path proses yang spesifik yang dapat diikuti pada saat runtime. Contoh urutan aktifitas.

Data mining juga dapat digunakan dalam jurnal ini kita mengeksplor teknik menggunakan data mining untuk memprediksikan path proses yang dapat diikuti saat runtime. Algoritma data mining dapat dipelajari di J48 C45 Naïve Bayes yang berjudul Sequential Minimal Optimisation and Multihost.

3.3. Kesimpulan

Business Process Management Systems, workflow systems, workflows, web proses dan web service menyediakan infrastruktur teknologi yang efisien dalam mendiniskan mengelola dan mendukung bisnis proses. Data degenerate dan dijalankan pada proses manajemen dimana biasanya disimpan dalam log system, suatu nilai dapat digunakan dan di ekstrak seperti eksekusi, lingkungan dan struktur.

Operasional organisasi di pasar modern ini seperti halnya kegiatan e-commerce , dan interaksi layanan distribusi web memerlukan QoS management.

References

Aalst, W.M.P.v.d., Dongen, B.F.v. *et al.* (2003) 'Workflow mining: a survey of issues and approaches', *Data and Knowledge Engineering (Elsevier)*, Vol. 47, No. 2, pp.237–267.

Agrawal, R., Gunopulos, D. *et al.* (1998) *Mining Process Models from Workflow Logs*, Sixth International Conference on Extending Database Technology, Valencia, Spain, Springer.

Anyanwu, K., Sheth, A. *et al.* (2003) 'Healthcare enterprise process development and integration', *Journal of Research and Practice in Information Technology, Special Issue in Health Knowledge Management*, Vol. 35, No. 2, pp.83–98.

CAPA (1997) *Course Approval Process Automation (CAPA)*, LSDIS Lab, Department of Computer Science, University of Georgia, Athens, GA.

Cardoso, J. (2002) *Stochastic Workflow Reduction Algorithm*, LSDIS Lab, Department of Computer Science, University of Georgia, Georgia, USA.

Cardoso, J. (2005) 'Path mining in web processes using profiles', *Encyclopedia of Data Warehousing and Mining*, Wang, J., Idea Group Inc., Pennsylvania, USA, pp.896–901.

Cardoso, J. and Sheth, A. (2003) 'Semantic e-workflow composition', *Journal of Intelligent Information Systems (JIIS)*, Vol. 21, No. 3, pp.191–225.

Cardoso, J., Bussler, C. *et al.* (2002) *Tutorial: Semantic Web Services and Processes: Semantic Composition and Quality of Service*, International Federated Conferences: DOA/ODBASE/CooPIS 2002, Irvine, CA.

Cardoso, J., Miller, J. *et al.* (2004) 'Modeling quality of service for workflows and web service processes', *Web Semantics: Science, Services and Agents on the World Wide Web Journal*, Vol. 1, No. 3, pp.281–308.

Chandrasekaran, S., Silver, G. *et al.* (2002) 'Service technologies and their synergy with simulation', *Proceedings of the 2002 Winter Simulation Conference (WSC'02)*, San Diego, California, pp.606–615.

Cortes, C. and Vapnik, V. (1995) 'Support vector networks', *Machine Learning*, Vol. 20, pp.273–297.

Cowley, S. (2002) *Study: BPM Market Primed for Growth*, September 23, http://www.infoworld.com/article/02/09/23/020923hnbpm_1.html.

Dietterich, T.G. (2000) 'Ensemble methods in machine learning', *First International Workshop on Multiple Classifier Systems*, Cagliari, Italy, Springer Verlag, New York.

Freund, Y. and Schapire, R. (1999) 'A short introduction to boosting', *Journal Japan Soc. for Artificial Intelligence*, Vol. 14, No. 5, pp.771–780.

Freund, Y. and Schapire, R.E. (1996) 'Experiments with a new boosting algorithm', *13th*

International Conference on Machine Learning, Morgan Kaufmann, 3–6 July, Bari, Italy, ISBN 1-55860-419-7.

Gillmann, M., Weikum, G. *et al.* (2002) 'Workflow management with service quality guarantees', *ACM SIGMOD'2002 International Conference on Management of Data*, Madison, Wisconsin.

Goldberg, D.E. (1989) *Genetic Algorithms in Search, Optimisation, and Machine Learning*, Addison Wesley, Reading, MA, USA.

Grigoria, D., Casatib, F. *et al.* (2004) 'Business process intelligence', *Computers in Industry*, Vol. 53, pp.321–343.

Hall, R.D., Miller, J.A. *et al.* (2003) 'Using workflow to build an information management system for a geographically distributed genome sequence initiative', in Prade, R.A. and Bohnert, H.J. (Eds.): *Genomics of Plants and Fungi*, Marcel Dekker Inc., New York, pp.359–371.

Herbst, J. and Karagiannis, D. (1998) 'Integrating machine learning and workflow management to support acquisition and adaption of workflow models', *Ninth International Workshop on Database and Expert Systems Applications*, Vienna, Austria, pp.745–752.

Kang, M.H., Froscher, J.N. *et al.* (1999) 'A multilevel secure workflow management system', *Proceedings of the 11th Conference on Advanced Information Systems Engineering*, Springer-Verlag, Heidelberg, Germany, pp.271–285.

Lee, C.H. and Shin, D.G. (1999) 'A multistrategy approach to classification learning in databases', *Data and Knowledge Engineering*, Vol. 31, No. 1, pp.67–93.

Lenic, M. and Kokol, P. (2002) 'Combining classifiers with Multimethod approach', *Second International Conference on Hybrid Intelligent Systems, Soft Computing Systems: Design, Management and Applications, Frontiers in Artificial Intelligence and Applications*, IOS Press, Amsterdam.

Lenic, M., Povalej, P. *et al.* (2003) 'Improved medical decision support with Multimethod approach', *IEEE Symposium on Computer-Based Medical Systems (CBMS 2003)*, IEEE Computer Society, New York, USA, pp.242–247.