

Magdeburger Schriften zur Wirtschaftsinformatik

Gamal Kassem

**Application Usage Mining:
Grundlagen und Verfahren**

Shaker Verlag
Aachen 2007

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Zugl.: Magdeburg, Univ., Diss., 2006

Copyright Shaker Verlag 2007

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8322-5995-2

ISSN 1618-2308

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen

Telefon: 02407/95 96 - 0 • Telefax: 02407/95 96 - 9

Internet: www.shaker.de • E-Mail: info@shaker.de

Vorwort

*„Daher ist die Aufgabe, nicht sowohl zu sehen,
was noch keiner gesehen hat, als bei
dem, was jeder sieht, zu denken, was noch
keiner gedacht hat“.*
(Schopenhauer)

Die vorliegende Arbeit entstand in der Arbeitsgruppe Wirtschaftsinformatik des Instituts für Technische und Betriebliche Informationssysteme der Fakultät für Informatik an der Otto-von-Guericke-Universität Magdeburg.

Für die umfangreiche Hilfe und Unterstützung während meiner Promotion habe ich vielen zu danken.

Meine Dankbarkeit gilt an erster Stelle meinem verehrten akademischen Lehrer und Doktorvater Herrn Prof. Dr. rer. pol. Claus Rautenstrauch für seine uneingeschränkte fachliche und persönliche Unterstützung. Seine zahlreichen inhaltlichen Anregungen und Fachgespräche und vor allem seine detaillierten fachlichen Korrekturen und Kommentare haben mich immer wieder auf die richtigen Pfade geleitet. Sein großes persönliches Engagement ist für einen Doktorvater keinesfalls selbstverständlich und weiß ich sehr zu schätzen.

Zu tiefem Dank verpflichtet bin ich meiner Familie. Meine Frau Iris Hollstein unterstützte und motivierte mich stets und übernahm die unangenehme Aufgabe des Korrektur-Lesens. Meine Tochter Mona übte sich in schier unendlicher Geduld, auch wenn es ihr oft schwer fiel. Mein Bruder Mohamed Kassem ermöglichte mir den Studienaufenthalt in Deutschland.

Weiterhin bedanke ich mich bei allen Mitarbeiterinnen und Mitarbeitern der Arbeitsgruppe sowie meinen Freunden, insbesondere Prof. Dr.-Ing. Jorge Marx Gomez, Jubran Rajub, Edeltraud Rautenstrauch, Ahmad Ghoneim, René Schultz, Dirk Dreschel und Christian Otto für ihre zahlreichen Anregungen und die fachliche und persönliche Unterstützung. Dem technischen Personal und ganz besonders Herrn Gerd Lange bin ich sehr dankbar für ihre Hilfe.

Ich widme diese Arbeit meine Mutter und der Seele meines Vaters.

Gamal Kassem

Im Februar 2007

Geleitwort

Wer kennt nicht folgende oder eine ähnliche Situation: Man trifft nach einer langen und strapaziösen Reise in seinem Hotel ein, nennt mit erwartungsfroher Miene seinen Namen, weist darauf hin, dass man reserviert hat, die anstrengend freundliche Dame an der Rezeption tippt Ihren Namen ein, noch mal, ... noch mal, ihre Wangen färben sich leicht rot, ... noch röter, und dann eröffnet sie die Situation mit folgendem Text: „Sie sind leider nicht im Computer“. Auch wenn man beiseite lässt, dass es physisch kaum möglich ist, einen erwachsenen Menschen in einen Computer zu stecken, zeigt sich hier ein alltägliches Dilemma: Die Abbildung der Realwelt auf die künstliche Welt des Computers gelingt nicht immer oder etwas radikaler formuliert: Man muss schon ganz schön naiv sein zu glauben, dass Systeme so genutzt werden, wie der Entwickler sich das vorgestellt hat.

Das Beispiel des enttäuschten Hotelgastes, der nun möglicherweise ein wenig anheimelndes Plätzchen unter der Brücke suchen muss, zeigt, dass solche Diskrepanzen zwischen realer und künstlicher Welt durchaus dramatische Folgen haben können. Während man den Vorfall aus dem Beispiel aber noch zu einem tragischen Einzelschicksal mit begrenztem Ausmaß herunterspielen kann, sind Unterschiede zwischen realweltlichen Sachverhalten und deren Abbildung in Informationssystemen in strategisch bedeutsamen und komplexen betrieblichen Anwendungen wie ERP-Systemen (ERP = Enterprise Resource Planning) für Unternehmen negativ erfolgswirksam, ohne dass die Ursache für Probleme bislang wirklich transparent wird.

Es macht daher durchaus Sinn, der Frage nachzugehen, was Benutzer wirklich mit den ihnen anvertrauten Informationssystemen machen. Die Beantwortung dieser Frage berührt allerdings sensible Bereiche: Berechtigte Interessen der Sicherung von Persönlichkeitsschutz und Privatheit dürfen nicht angetastet werden. Elektronische Spionage ist demnach schon aus ethischen, aber auch aus rechtlichen Gründen kein probater Lösungsansatz.

Gamal Kassem legt mit diesem Werk die Grundlage für die Methodik des Application Usage Mining, welches die Rekonstruktion von tatsächlich durchgeführten Workflows aus den Trace-Files von ERP-Systemen erlaubt. Diese Methodik ist auf der einen Seite hinreichend anonym, sodass Anforderungen des Daten- und Persönlichkeitsschutzes uneingeschränkt erfüllt werden, auf der anderen Seite aber auch hinreichend präzise, sodass Schwachstellen in der Workflow-Ausführung tatsächlich identifiziert und beseitigt werden können. Dabei muss die Schwachstellenbeseitigung der nicht zwangsläufig zu einer Maßregelung der Benutzer führen, da die Ursachen auch in einer realitätsfremden Workflow-Modellierung oder fehlerhaftem Customizing liegen können.

Die vorliegende Arbeit liefert damit einen wertvollen Beitrag, Anwendungssysteme für den Anwender besser anwendbar zu gestalten. Es ist zu hoffen, dass sie einen angemessenen Leserkreis findet.

Claus Rautenstrauch

Im Februar 2007

Inhaltsverzeichnis

Vorwort	I
Geleitwort	III
Inhaltsverzeichnis	V
Abkürzungsverzeichnis	VIII
Abbildungsverzeichnis	IX
Tabellenverzeichnis	XI
Symbolverzeichnis	XII
1 Einleitung	1
1.1 Ausgangssituation und Problemstellung	1
1.2 Ziel	3
1.3 Aufbau	3
2 Grundlagen und Begriffserklärungen	5
2.1 Geschäftsprozesse	5
2.2 Workflow	5
2.3 Petri-Netze	8
2.3.1 Netze und ihre Strukturen	9
2.3.2 Stellen/Transitionen-Netze	10
2.3.3 Markierte S/T-Netze	11
2.4 ERP-Systeme	13
3 Trace-, Log- und Protokoll-Daten als Grundlage für Mining-Verfahren	15
3.1 Web Usage Mining	16
3.2 Process Mining	21
3.3 Workflow Mining	24
3.4 Problem des Mining im ERP-Systembereich	28
3.5 Der Application Usage Mining Prozess	30
3.6 Trace-, Log- und Protokoll-Daten in ERP-Systemen	32
4 Das ERP-System am Beispiel von SAP ECC	34
4.1 Die Architektur eines SAP-Systems	34
4.2 Technische Realisierung von Geschäftsprozessen	36
4.2.1 SAP-Transaktionen	37
4.2.2 Dialogorientierte Anwendungen in SAP ECC	37
4.3 Die Nutzerinteraktion mit der Dialoganwendung	39
4.4 Trace-Dateien in SAP ECC	43
4.4.1 System-Traces	43
4.4.2 Entwickler-Traces	44
5 Vorbereitungsphase des AUM	47
5.1 Die Integration von Interaktionsdaten	47
5.1.1 SAP-Transaktionen	50

5.1.2	Funktionen	52
5.1.3	Nachricht	54
5.1.4	Workflow	56
5.1.5	Interaktionsschritt.....	57
5.2	Extraktion von Meta-Interaktionsdaten.....	58
5.2.1	ECC-Repository	59
5.2.1.1	Business-Engineering-Objekte	60
5.2.1.2	Programmobjekte.....	63
5.2.1.3	Dictionary-Objekte	65
5.2.2	Benutzerinformationssystem.....	66
5.2.3	Workflow-Beschreibungen	67
5.3	Extraktion von technischen Interaktionsdaten	67
6	Die Muster-Entdeckungsphase - Ermittlung von Workflow-Instanzen	71
6.1	Lösungs-Ansatz zum Workflow-Instanz-Problem.....	71
6.2	Einleitung von funktionsbezogenen Daten als Function-Trace-File aus dem Daten-Pool.....	72
6.3	Function-Trace-File.....	74
6.4	Entdeckung von Workflow-Instanzen im Function-Trace-File	74
6.5	Behandlung von Konfliktfällen	85
6.6	Entdeckung von Aufgabenschritten in ERP-Systemen als kleinster Detaillierungsgrad eines Workflows.....	99
6.6.1	Detaillierungsgrad eines Workflows.....	99
6.6.2	Ermittlung von Funktionsschritt-Instanzen aus dem Function-Trace-File	102
7	Die Muster-Analyse-Phase	104
7.1	Workflow-Netz.....	104
7.2	Das Routing von Workflows.....	108
7.2.1	Sequenzielles Routing.....	108
7.2.2	Paralleles Routing	109
7.2.3	Verzweigtes Routing.....	109
7.2.4	Darstellung von Verknüpfungsoperatoren	111
7.2.5	Routing von Iteration	112
7.3	Die α -Algorithmen	113
7.4	Der Workflow-Fall-Generator.....	116
7.5	Der Generierungsprozess des Workflow-Fall-Generators	117
7.6	Beispiel zu Generierung von Workflow-Modellen	119
8	Integration von AUM in den Continuous-Improvement-Prozess.....	129
8.1	Business Process Reengineering und Continuous Improvement Process	129
8.2	Mikro- und Makroprozesse	130
8.3	Phasen des Continuous Improvement	131
8.4	Integration der AUM-Phasen in den Workflow-Verbesserungslebenslauf.....	133

8.5	Verbesserungsmaßnahmen in der Modellierungs- und Implementierungsphase des Workflow-Verbesserungslebenslaufs	135
9	Schlussbetrachtung	137
9.1	Zusammenfassung	137
9.2	Ergebnisse der Arbeit	138
9.3	Technische Probleme bei der Realisierung von AUM	138
9.4	Weitere mögliche Einsatz-Gebiete des AUM	139
9.5	Aussicht	140
	Anhang A: Zugriffpfad-File	142
	Anhang B: Function-Trace-File	148
	Literaturverzeichnis	150