

Editorial

Bernhard Mitschang · Daniela Nicklas · Theo Härder

Online publiziert: 12. Juni 2015
© Springer-Verlag Berlin Heidelberg 2015

Schwerpunktthema: Data Management for Mobility

Mobilität – Bewegung – erzeugt heutzutage nicht nur Wärme in Form von Reibung oder chemischen Verbrennungsprozessen, sondern auch Daten in vielfältiger Form: Kameras, Radarsysteme und Induktionsschleifen messen Verkehrsströme, Reisende fotografieren, twittern und bewerten beliebte Orte, und Kunden-, Eintritts- und Kreditkarten erzeugen bei jeder Transaktion eine Datenspur, mit der sich die Bewegung der Besitzerin detailliert verfolgen lässt. Auch Logistiksysteme, Mobilfunkprovider und intelligente Produktionsanlagen sammeln inzwischen große Mengen an Daten, denen die Bewegung und Zustandsänderung von Gütern und Menschen zugrunde liegt.

Für das Datenmanagement ergeben sich daraus Herausforderungen, die sich grob in zwei Forschungsrichtungen klassifizieren lassen: die effiziente und skalierbare Verwaltung von *Moving Objects*, also beweglichen Objekten, wie Fahrzeugen, Personen oder Gütern, sowie *Mobile Computing*, sprich die Entwicklung von Informatiksystemen mit sich bewegenden Komponenten, wie z. B. mobilen, ortsbezogenen Anwendungen oder Fahrerassistenzsystemen. In bei-

den Forschungsrichtungen finden sich zahlreiche Herausforderungen an das Datenmanagement, denen sich die internationale Forschungsgemeinschaft wie auch die Industrie seit einigen Jahren widmet.

In diesem Themenheft sind drei Beiträge enthalten, die aktuelle Herausforderungen an das Datenmanagement für Mobilität adressieren. Der erste Beitrag, „Optimized P2P Data Management for Reliable Workflow Execution in Mobile Environments“ von Nenad Stojnic und Heiko Schuldt (Universität Basel) stammt aus dem Bereich *Mobile Computing*: es wird gezeigt, wie Workflows auf einer Systemumgebung zuverlässig ausgeführt werden können, deren Komponenten mobil sind und deswegen nicht dauerhaft zur Ausführung bereit stehen. Als Beispiel werden (stationäre) Sensornetze genannt, deren Daten mit mobilen Endgeräten ausgelesen und auch verarbeitet werden. Die Autoren stellen ein Workflow-System vor, das bestehende Ansätze zur effizienten Datenverwaltung in Peer-to-Peer-Systemen erweitert, so dass eine zuverlässige Ausführung von Workflows möglich wird.

Auch im zweiten Beitrag geht es um mobile Anwendungen, hier wird jedoch der Aspekt der Sicherheit adressiert. Christoph Stach und Bernhard Mitschang (Universität Stuttgart) zeigen auf, welche zusätzlichen Angriffsvektoren sich hier ergeben. Seit 2004, als der erste Wurm speziell für mobile Geräte aufkam, wächst die Anzahl an Schadsoftware kontinuierlich. Heute werden täglich bereits über 2000 neue Schadprogramme alleine für Android-basierte Geräte entdeckt. Der Grund hierfür ist einfach. Auf den mobilen Geräten ist eine Vielzahl an personenbezogenen Daten gespeichert (u. a. Positionsdaten oder Kontakte) und kombiniert lässt sich daraus ein detailliertes Nutzerprofil erstellen. Neben kriminellen Organisationen setzen daher auch beispielsweise Firmen, die das Kundenverhalten analysieren wollen, um so personalisierte Werbung schalten zu können,

T. Härder (✉)
AG Datenbanken und Informationssysteme, TU Kaiserslautern,
67663 Kaiserslautern, Deutschland
E-Mail: haerder@cs.uni-kl.de

D. Nicklas
Lehrstuhl für Informatik, insb. Mobile Systeme/Mobilität,
Otto-Friedrich-Universität Bamberg, An der Weberei 5, 96047
Bamberg, Deutschland
E-Mail: daniela.nicklas@uni-bamberg.de

B. Mitschang
Institut für Parallele und Verteilte Systeme, Universität Stuttgart,
Universitätsstraße 38, 70569 Stuttgart, Deutschland
E-Mail: bernhard.mitschang@ipvs.uni-stuttgart.de

oder selbst Geheimdienste verstärkt Spionagesoftware ein. Aus diesem Grund beschäftigen sich viele Forschungsprojekte damit, wie sich private Daten auf Mobilplattformen besser schützen lassen. Es zeigt sich allerdings, dass diese Lösungen meist nur wenige Aspekte eines Datensicherheitsystems (u. a. Vertraulichkeit, Nachweisbarkeit oder Privatheit) adressieren. Auch hinsichtlich der Nutzerfreundlichkeit lassen diese Lösungen oft zu wünschen übrig, da sie sich nur grobgranular konfigurieren lassen, sie sich nicht auf Datenquellen anwenden lassen und bei einem Berechtigungsentzug zu einem Absturz führen. Im Beitrag wird die „Privacy Management Platform“ durch den neu eingeführten „Secure Data Container“ zu einem ganzheitlichen Datensicherheitsystem erweitert. Durch diese Kombination werden nicht nur alle Schutzziele erfüllt, sondern es werden auch hinsichtlich Performanz sehr gute Ergebnisse erzielt.

Der dritte Beitrag widmet sich schließlich einer Technologie, die sich für die kontinuierliche Verarbeitung von Daten, wie sie auch in mobilen Umgebungen entstehen, oft besser eignet als eine persistente, ständig wachsende Datenspeicherung. Beim sogenannten Complex Event Processing werden zeitlich geordnete Datensätze als Ereignisse (Events) interpretiert, die durch verschiedene Ableitungen und Muster zu komplexeren Ereignissen (Complex Events) verarbeitet werden. Dabei erfolgt in der Regel eine Interpretation, die aus Rohdaten Information und Wissen gewinnt, ein für viele Anwendungsgebiete notwendiger Prozess. Im Beitrag „Complex Event Processing on Linked Stream Data“ zeigen Omran Saleh, Stefan Hagedorn und Kai-Uwe Sattler (TU Ilmenau), wie sich diese Technologie auf semantisch annotierten Datenströmen (RDF) anwenden lässt und stellen dafür ihr PipeFlow-System vor, ein relationales Datenflusssystem, das gezielt durch RDF-Operatoren erweitert wurde.

Fachbeiträge in diesem Heft

Die drei Beiträge zum Schwerpunktthema dieses Heftes werden durch zwei Fachbeiträge ergänzt. Im Beitrag *Toward GPU-accelerated Database Optimization* beschreiben Andreas Meister (Universität Magdeburg), Sebastian Breß (TU Dortmund) und Gunter Saake (Universität Magdeburg) verbesserte Ansätze zur Anfrageoptimierung durch Einsatz von Co-Prozessoren wie beispielsweise GPUs. Bisher basieren die meisten Optimierungsansätze auf sequentiellen Algorithmen, deren Parallelisierung das Leistungspotenzial heutiger und künftiger Rechensysteme wesentlich besser ausnutzen würde. Da zudem die Anfrageoptimierung eine enorm rechenintensive Aufgabe ist, muss auf eine kosteneffektive Ausgleichsstrategie zwischen Anfrageverarbeitung und ihrer Optimierung auf CPUs und GPUs geachtet werden. In ihrem Beitrag motivieren die Autoren den Einsatz von parallelen Co-Prozessoren für die DB-Optimierung und identifizieren Optimierungsprobleme, die von einer Parallelisierung

profitieren können. Schließlich entwerfen sie Optimierungsheuristiken zur parallelen Platzierung von Anfrageoperatoren bei hybriden CPU/Co-Prozessor-Systemen und spezialisieren sie auf eine GPU-Nutzung.

Der Beitrag *Managed Query Processing within the SAP HANA Database Platform* von Norman May, Alexander Böhm, Meinolf Block (SAP SE, Deutschland) und Wolfgang Lehner (TU Dresden) zeigt, wie SAP HANA die Einsatzmöglichkeiten von herkömmlichen Datenbank-Servern erweitert und dadurch Datenmodelle, die auf Texten, Graphen oder Hierarchien beruhen, unterstützen kann. Die Autoren geben einen ganzheitlichen Überblick über die HANA-Plattform – von der Datenmodellierung über die Softwareverteilung im Gesamtsystem bis zur effizienten (web-basierten) Anwendungsausführung – und skizzieren, wie ihr spezielles Leistungsvermögen genutzt werden kann. Als wesentliches Unterscheidungsmerkmal integriert die HANA-Plattform die meisten Schritte im „Lifecycle“ einer Anwendung und kann so sicherstellen, dass alle relevanten Komponenten konsistent bleiben, wann immer sie modifiziert werden. Weiterhin gehören zur HANA-Plattform auch Transport- und Kommunikationsmöglichkeiten, um Anwendungen in einer komplexen Systemlandschaft zu laden und zu entladen.

Community-Beiträge in diesem Heft

In der Rubrik „Datenbankgruppen vorgestellt“ finden Sie unter *Der Lehrstuhl Datenbank- und Informationssysteme der Universität Rostock (LS DBIS)* einen Beitrag von Andreas Heuer, Holger Meyer, Ilvio Brunder und Meike Klettke (Universität Rostock). Im Jahr 2014 feierte dieser LS DBIS sein zwanzigjähriges Bestehen. Zur Jubiläumsveranstaltung mit ehemaligen und aktuellen Studenten, Mitarbeitern, Kollegen und Kooperationspartnern wurde diverses Material aus 20 Jahren aufbereitet (<http://www.ls-dbis.de/20jahre>). Daran geben die Autoren in diesem Beitrag einen interessanten Rückblick über die Forschung und Lehre im Bereich Datenbank- und Informationssysteme, die in den 90er und 2000er Jahren in Rostock geleistet wurde. Weiterhin erläutern sie drei Langzeitvorhaben, die einen Ein- und Ausblick auf aktuelle Forschungsarbeiten bieten. Schließlich skizzieren sie unter dem Stichwort „Bologna invers“ die Datenbanklehre an der Universität Rostock und wie sie es geschafft haben, das alte Konzept von studentischen Projektgruppen trotz Bologna-Überregulierung als sinnvolles Studienelement wieder aufleben zu lassen.

Weiterhin erscheint in der Rubrik „Community“ mit dem Beitrag „BTW 2015 – Jubiläum an der Waterkant“ von Wolfram Wingerath, Steffen Friedrich und Norbert Ritter (Universität Hamburg) ein ausführlicher Bericht über die 16. GI-Fachtagung „Datenbanksysteme für Business, Technologie

und Web“, die Anfang März 2015 an der Universität Hamburg stattfand.

In diesem Heft bietet die Rubrik „Dissertationen“ mit neun Kurzfassungen von Dissertationen aus der deutschsprachigen DBIS-Community wieder einen erfreulich großen Umfang.

Die Rubrik „Community“ enthält schließlich unter *News* weitere aktuelle Informationen aus der DBIS-Gemeinde.

Künftige Schwerpunktthemen

1 Best Workshop Papers of BTW 2015

This special issue of the „Datenbank-Spektrum“ is dedicated to the Best Papers of the Workshops running at the BTW 2015 at the University of Hamburg. The selected Workshop contributions should be extended to match the format of regular DASP papers.

Issue delivery: DASP-3-2015 (November 2015)

Guest editor:

Theo Härder, University of Kaiserslautern
haerder@cs.uni-kl.de

2 Big Data & IR

The term *Big Data* refers to data and respective processing strategies, which, due to their sheer size, require a data center for the processing, and which become available through the ubiquitous computer and sensor technology in many facets of everyday life. Interesting scientific questions in this regard are the organization and management of Big Data, but also the identification of problems that now can be studied and better understood through the collection and analysis of Big Data. In the context of information retrieval as the purposeful search for relevant content, there are two main challenges: 1) retrieval in Big Data and 2) improved retrieval because of Big Data.

Retrieval in Big Data focuses on the organization, the management, and the quick access to Big Data, but also addresses the creative process of identifying interesting research questions that can only be understood and answered in Big Data. Besides the development of powerful frameworks for the maintenance and analysis of text, multimedia, sensor, and simulation data, an important research direction is the question of what kind of insights Big Data may give us today and in the future.

The second challenge in the context of Big Data & IR is the improvement of retrieval approaches through Big Data.

Examples include the classic question of improved Web or eCommerce search via machine learning on user behavior data, the usage of user context for retrieval, or the exploitation of semantic data like Linked Open Data or knowledge graphs.

We are looking for contributions from researchers and practitioners in the above described context. The contributions may be submitted in German or in English and should observe a length of 8–10 pages in the Datenbank-Spektrum format (cf. the author guidelines at www.datenbank-spektrum.de).

Important dates:

- Notice of intent for a contribution: August 15th, 2015
- Deadline for submissions: October 1st, 2015
- Issue delivery: DASP-1-2016 (March 2016)

Guest editors:

Matthias Hagen, Universität Weimar
matthias.hagen@uni-weimar.de
Benno Stein, Universität Weimar
benno.stein@uni-weimar.de

3 Schutz der Privatsphäre in einer ubiquitären Welt

Mit immer mehr mobilen Geräten und Sensoren werden u. a. große Mengen an persönlichen Daten gesammelt, verarbeitet und transformiert. Solche Sammlungen personenbezogener Daten sind auf der einen Seite notwendig, um personenspezifische Angebote machen zu können, die dem Empfänger örtlich und zeitlich von Nutzen sind, oder um Trends zu erkennen und somit Planungen in unterschiedlichen Bereichen genauer und effizienter ausführen zu können. Auf der anderen Seite dienen sie häufig dazu, individuelle Personenprofile zu erstellen, die zum Vorteil oder Nachteil der beschriebenen Person genutzt werden können.

Aus den genannten Gründen wird es immer wichtiger, den Datenschutz in einer ubiquitären Welt im Kontext von Big Data nicht nur juristisch abzusichern (Bundesdatenschutzgesetz). Vielmehr wird es immer dringlicher, auch technische Möglichkeiten, Mechanismen und Ansätze zu entwerfen und zu realisieren, die es Personen ermöglichen, die Kontrolle über ihre Daten besser spezifizieren sowie ihre Nutzung und Weitergabe besser kontrollieren und nachvollziehen zu können. Trotz großer Fortschritte im Bereich des Schutzes der Privatsphäre durch unterschiedliche Techniken besteht weiterhin eine große Herausforderung darin, skalierbare Ansätze und Lösungen sowohl für die Nutzung personenbezogener Daten durch Dritte als auch deren Kontrolle durch den „Spender“ zu entwickeln und zu realisieren.

Somit ist es das Ziel des Themenheftes, neben einer Einführung in das Thema skalierbare Ansätze und Lösungen für das Sammeln, Verarbeiten und Analysieren personenbezogener Daten in unterschiedlichen Anwendungsdomänen zu beschreiben. Mögliche Themen für dieses Themenheft sind (nicht ausschließlich):

- Schutz der Privatsphäre im Bereich Big Data generell beim Sammeln, Integrieren, Verarbeiten und Analysieren von Daten
- Technische Umsetzung juristischer (gesetzlicher) Vorgaben entsprechend des deutschen Rechts bzw. des EU-Rechts und Vorgaben in unterschiedlichen Bereichen wie beispielsweise dem Gesundheitsbereich (Arzt, Krankenhaus, Versicherer) oder dem Finanzbereich (Banken)
- Sprachen zur Beschreibung von Privacy-Präferenzen und deren Überprüfung (in skalierbarer Form)
- Datenaustausch unter Berücksichtigung von Privacy-Präferenzen und gesetzlichen Vorgaben
- Anfragebearbeitung in Datenbanksystemen unter Berücksichtigung von Privacy-Präferenzen
- Quantitative Bewertung von Ansätzen zum Schutz der Privatsphäre im Kontext der Nutzung geschützter Daten
- Anforderungsanalysen für den Schutz der Privatsphäre in Anwendungsdomänen – Schutz gegen Genauigkeit der Daten
- Anforderungen des Schutzes der Privatsphäre für räumliche und zeitbezogene Daten in verschiedenen Anwendungsbereichen
- Infrastrukturen zum Schutz der Privatsphäre
- Modelle zum Schutz der Privatsphäre bei Zugriff oder Datennutzung

Beitragsformat: 8–10 Seiten, zweispaltig

Ankündigung einer Beitragseinreichung bis zum 15. Dezember 2015

Gastherausgeber:

Johann-Christoph Freytag, HU Berlin

freytag@informatik.hu-berlin.de

Eric Buchmann, Karlsruher Institut für Technologie

eric.buchmann@kit.edu

Einreichung der Beiträge für DASP-2-2016 bis zum 1. Februar 2016

4 Data Management for Bio- and Geosciences

Like many other scientific disciplines, research in the bio- and geosciences follow more and more a data-driven approach. Big Data in the classical sense is only one of the issues, but probably more often the everyday problem of scientists is to cope with lots of „small“, heterogeneous data that needs to be integrated to answer complex questions.

This special issue addresses the arising challenges and solutions for data management in these areas. We are interested in both survey papers and papers describing original research dealing with the following or similar topics in the context of the bio- and geosciences:

- Data-intensive science
- Data management
- Data integration
- Spatio-temporal data processing
- Scientific workflows
- Semantic web technologies
- Visualization and visual analytics
- Data stream management
- Case studies and applications

Important dates:

- Notice of intent for a contribution: April 15th, 2016
- Deadline for submissions: June 1st, 2016
- Issue delivery: DASP-3-2016 (November 2016)

Paper format: 8–10 pages, double column

Guest editors:

Bernhard Seeger, Philipps-Universität Marburg

seeger@mathematik.uni-marburg.de

Birgitta König-Ries, Friedrich-Schiller-Universität Jena

Brigitta.Koenig-Ries@uni-jena.de