A Two-Tiered Segmentation Approach for Transaction Data Warehousing

Data warehousing populates data from different source systems into a central data warehouse (DW) through extraction, transformation, and loading (ETL). Massive transaction data are routinely recorded in a variety of applications such as retail commerce, bank systems, and website management. Transaction data record the timestamp and relevant reference data needed for a particular transaction record. It is a non-trivial task for a standard ETL to process transaction data with dependencies and high velocity. This chapter presents a two-tiered segmentation approach for transaction data warehousing. The approach uses a so-called two-staging ETL method to process detailed records from operational systems, followed by a dimensional data process to populate the data store with a star or snowflake schema. The proposed approach is an all-in-one solution capable of processing fast/slowly changing data and early/late-arriving data. This chapter evaluates the proposed method, and the results have validated the effectiveness of the proposed approach for processing transaction data.
Big Data Technology-Enabled Analytical Solution for Quality Assessment of Higher Education Systems

Educational intelligence is a broad area of big data analytical applications that make use of big data technologies for implementation of solutions for education and research. This paper demonstrates the designing, development and deployment of an educational intelligence application for real-world scenarios. Firstly, a quality assessment framework for higher education systems that evaluate institutions on the basis of performance of outgoing students was proposed. Secondly, big data enabled technological setup was used for its implementation. Literature was surveyed to evaluate existing quality frameworks. Most existing quality assessment systems take into account the dimensions related to inputs, processes and outputs, but they tend to ignore the perspective that assesses the institution on the basis of outcome of the educational process. This paper demonstrates the use of outcome perspective to compute quality metrics and create visual analytics. In order to implement and test the framework, R programming language and a cloud based big data technology that is Google, BigQuery were used.

General information
Publication status: Published
Organisations: Sustainability, Energy Systems Analysis, Department of Technology, Management and Economics, Jamia Millia Islamia, Princess Nourah Bint Abdulrahman University
Contributors: Khan, S., Alam, M., Liu, X., Shakil, K. A.
Pages: 292-304
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: International Journal of Advanced Computer Science and Applications
Volume: 10
Issue number: 6
ISSN (Print): 2156-5570
Original language: English
Research output: Contribution to journal › Journal article – Annual report year: 2019 › Research › peer-review

Statistical Analysis for Factors Influencing Electricity Consumption at Regional Level

A regional analysis of factors influencing electricity consumption is performed using panel dataset. It was identified that climate factors and an economic factor of hotel occupancy could be used to characterise the regional monthly electricity consumption. This study selects Jakarta, Indonesia as the case study, covering the monthly time period from 2006-2012. Multiple linear regression approach is applied to analyse the association between the continues variables. The computation and visualisation is executed in R. One of the explanatory variables, namely hotel room occupancy, is interesting to be analysed in this case because the case study is the capital of the country. Many national events are being held centrally in Jakarta. In practice, limited studies have used this variable to model electricity consumption. The finding shows that the simplest model is conducted by excluding the variable day of rain and variable rainfall that slightly contribute to the model. The result identifies that the relationship between electricity consumption and hotel occupancy is statistically significant, as well as the relationship between electricity consumption and temperature.

General information
Publication status: Published
Organisations: Department of Technology, Management and Economics, Sustainability, Energy Systems Analysis, Sam Ratulangi University
Contributors: Kewo, A., Manembu, P., Liu, X., Nielsen, P. S.
Number of pages: 6
Pages: 132-137
Publication date: 2019

Host publication information
Title of host publication: 2018 IEEE 7th International Conference on Power and Energy (PECon)
Publisher: Institute of Electrical and Electronics Engineers Inc.
Article number: 8684037
ISBN (Electronic): 9781538654316
Keywords: Electricity consumption, modelling, multiple linear regression, statistical analysis
DOIs: 10.1109/PECON.2018.8684037
Source: Scopus
Source ID: 85064757470
Research output: Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report year: 2019 › Research › peer-review
Two approaches for synthesizing scalable residential energy consumption data

Many fields require scalable and detailed energy consumption data for different study purposes. However, due to privacy issues, it is often difficult to obtain sufficiently large datasets. This paper proposes two different methods for synthesizing fine-grained energy consumption data for residential households, namely a regression-based method and a probability-based method. They each use a supervised machine learning method, which trains models with a relatively small real-world dataset and then generates large-scale time series based on the models. This paper describes the two methods in detail, including data generation process, optimization techniques, and parallel data generation. This paper evaluates the performance of the two methods, which compare the resulting consumption profiles with real-world data, including patterns, statistics, and parallel data generation in the cluster. The results demonstrate the effectiveness of the proposed methods and their efficiency in generating large-scale datasets.

General information
Publication status: Published
Organisations: Sustainability, Energy Systems Analysis, Department of Technology, Management and Economics, Department of Civil Engineering, Energy and Services, University College of Northern Denmark, University of Technology Sydney
Corresponding author: Liu, X.
Contributors: Liu, X., Ifitkhar, N., Huo, H., Li, R., Nielsen, P. S.
Pages: 586-600
Publication date: 2019
Peer-reviewed: Yes

Publication information
Journal: Future Generation Computer Systems
Volume: 96
ISSN (Print): 0167-739X
Ratings:
BFI (2019): BFI-level 1
Web of Science (2019): Indexed yes
Original language: English
Keywords: Energy consumption, Time series, Synthesize, Simulation, Data generation
DOI:
10.1016/j.future.2019.01.045
Source: FindIt
Source ID: 2443610223
Research output: Contribution to journal › Journal article – Annual report year: 2019 › Research › peer-review

Analysis and Visualization of Urban Emission Measurements in Smart Cities
Cities worldwide aim to reduce their greenhouse gas emissions and improve air quality for their citizens. Therefore, there is a need to implement smart city approaches to monitor, model, and understand local emissions to better guide these actions. We present our approach that deploys a number of low-cost sensors through a wireless Internet of Things (IoT) backbone and is thus capable of collecting high-granular data. Based on a flexible architecture, we built an ecosystem of data management and data analytics including processing, integration, analysis, and visualization as well as decision-support systems for cities to better understand their emissions. Our prototype system has so far been tested in two Scandinavian cities. We present this system and demonstrate how to collect, integrate, analyze, and visualize real-time air quality data.

General information
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, Norwegian University of Science and Technology, AIA Science
Number of pages: 4
Publication date: 2018

Host publication information
Title of host publication: Proceedings of the 21st International Conference on Extending Database Technology (EDBT)
Publisher: OpenProceedings.org
ISBN (Electronic): 978-3-89318-078-3
Electronic versions:
ctt_data_analysis_ecosystem_2_.pdf
URLs:
Architecture Design of Smart Meter Controlling System for Dynamic IP Environments

Smart meters are the electronic devices for measuring energy consumption in real time. Usually, static public IP (Internet Protocol) addresses are allocated to realize the point-to-point (P2P) communication and remote controlling for smart metering systems. This, however, restricts the wide deployment of smart meters, due to the deficiency of public IP resources. This paper describes the road map architecture development of smart meter controlling system with dynamic IP addresses. We start our development from the naive architecture to a novel subscription-based communication architecture (SCA) for the support of dynamic IP addresses and group controlling of smart meters. The paper evaluates the proposed architecture by comparing the traditional P2P architecture, and validate its effectiveness to interact with smart meters.

Clustering-based analysis for residential district heating data

The wide use of smart meters enables collection of a large amount of fine-granular time series, which can be used to improve the understanding of consumption behavior and used for consumption optimization. This paper presents a clustering-based knowledge discovery in databases method to analyze residential heating consumption data and evaluate information included in national building databases. The proposed method uses the K-means algorithm to segment consumption groups based on consumption intensity and representative patterns and ranks the groups according to daily consumption. This paper also examines the correlation between energy intensity and the characteristics of buildings and occupants, load profiles of households, consumption behavior changes over time, and consumption variability. The results show that the majority of the customers can be represented by fairly constant load profiles. Calendar context has an impact not only on the patterns but also on the consumption intensity and user behaviors. The variability studies show that consumption patterns are serially correlated, the customers with high energy consumption have lower variability, and the consumption is more stable over time. These findings will be valuable for district heating utilities and energy planners to optimize their operations, design demand-side management strategies, and develop targeting energy-efficiency programs or policies.
Recognizing Textual Entailment with Attentive Reading and Writing Operations

Inferencing the entailment relations between natural language sentence pairs is fundamental to artificial intelligence. Recently, there is a rising interest in modeling the task with neural attentive models. However, those existing models have a major limitation to keep track of the attention history because usually only one single vector is utilized to memorize the past attention information. We argue its importance based on our observation that the potential alignment clues are not always centralized. Instead, they may diverge substantially, which could cause the problem of long-range dependency. In this paper, we propose to facilitate the conventional attentive reading operations with two sophisticated writing operations - forget and update. Instead of utilizing a single vector that accommodates the attention history, we write the past attention information directly into the sentence representations. Therefore, higher memory capacity of attention history could be achieved. Experiments on Stanford Natural Language Inference corpus (SNLI) demonstrate the superior efficacy of our proposed architecture.

General information
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, Coventry University, University of Shanghai for Science and Technology
Pages: 847-860
Publication date: 2018

Host publication information
Title of host publication: Database Systems for Advanced Applications. DASFAA 2018.
Publisher: Springer
(Lecture Notes in Computer Science, Vol. 10827).
DOIs:
10.1007/978-3-319-91452-7_54
Source: PublicationPreSubmission
Source ID: 147928532
Research output: Chapter in Book/Report/Conference proceeding – Article in proceedings – Annual report year: 2018 – Research – peer-review

Scalable Prediction-based Online Anomaly Detection for Smart Meter Data

Abstract Today smart meters are widely used in the energy sector to record energy consumption in real time. Large amounts of smart meter data have been accumulated and used for diverse analysis purposes. Anomaly detection raises the big data problem, namely the detection of abnormal events or unusual consumption behaviors. However, there is a lack of appropriate online systems that can handle anomaly detection for large-scale smart meter data effectively and efficiently. This paper proposes a lambda system for detecting anomalous consumption patterns, aiming at assisting decision makings for smart energy management. The proposed system uses a prediction-based detection method, combined with a novel lambda architecture for iterative model updates and real-time anomaly detection. This paper evaluates the system using a real-world data set and a large synthetic data set, and compares with three baselines. The results show that the proposed system has good scalability, and has a competitive advantage over others in anomaly detection.

General information
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis
Corresponding author: Liu, X.
Contributors: Liu, X., Nielsen, P. S.
Number of pages: 14
Pages: 34-47
Publication date: 2018
Air Quality Monitoring System and Benchmarking

Air quality monitoring has become an integral part of smart city solutions. This paper presents an air quality monitoring system based on Internet of Things (IoT) technologies, and establishes a cloud-based platform to address the challenges related to IoT data management and processing capabilities, including data collection, storage, analysis, and visualization. In addition, this paper also benchmarks four state-of-the-art database systems to investigate the appropriate technologies for managing large-scale IoT datasets.

General information
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis
Contributors: Liu, X., Nielsen, P. S.
Pages: 459-470
Publication date: 2017

Host publication information
Title of host publication: International Conference on Big Data Analytics and Knowledge Discovery
Publisher: Springer
(Lecture Notes in Computer Science, Vol. 10440).
Electronic versions: shortpaper.pdf. Embargo ended: 01/10/2018
DOIs:
10.1007/978-3-319-64283-3_34
Source: FindIt
Source ID: 2372953457
Research output: Chapter in Book/Report/Conference proceeding – Article in proceedings – Annual report year: 2017 – Research – peer-review

A Novel Smart Meter Controlling System with Dynamic IP Addresses

Smart meters are the electronic devices for measuring energy consumption in real time. Usually, static public IP addresses are allocated to realize the point-to-point (P2P) communication and remote controlling for smart metering systems. This, however, restricts the wide deployment of smart meters, due to the deficiency of public IP resources. This paper proposes a novel subscription-based communication architecture for the support of dynamic IP addresses and group controlling of smart meters. The paper evaluates the proposed architecture by comparing the traditional P2P architecture, and validate its effectiveness to interact with smart meters.

General information
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, Sam Ratulangi University, De La Salle University-Manila
Pages: 1465-1470
Publication date: 2017

Host publication information
A Scalable Smart Meter Data Generator Using Spark

Today, smart meters are being used worldwide. As a matter of fact smart meters produce large volumes of data. Thus, it is important for smart meter data management and analytics systems to process petabytes of data. Benchmarking and testing of these systems require scalable data, however, it can be challenging to get large data sets due to privacy and/or data protection regulations. This paper presents a scalable smart meter data generator using Spark that can generate realistic data sets. The proposed data generator is based on a supervised machine learning method that can generate data of any size by using small data sets as seed. Moreover, the generator can preserve the characteristics of data with respect to consumption patterns and user groups. This paper evaluates the proposed data generator in a cluster based environment in order to validate its effectiveness and scalability.

A Science Cloud for Smart Cities Research

Cities are densely populated and heavily equipped areas with a high level of service provision. Smart cities can use these conditions to achieve the goals of a smart society for their citizens. To facilitate such developments, the necessary IT-infrastructure has to be in place for supporting, amongst many other things, the whole lifecycle of big data management and analytics for research activities. At the Centre for IT-Intelligent Smart Energy for Cities, we have therefore been developing a flexible infrastructure, based on open sourcetechnologies. This paper presents this solution and its application in a city and building research.
A Survey of Scholarly Data: From Big Data Perspective

Recently, there has been a shifting focus of organizations and governments towards digitization of academic and technical documents, adding a new facet to the concept of digital libraries. The volume, variety and velocity of this generated data, satisfies the big data definition, as a result of which, this scholarly reserve is popularly referred to as big scholarly data. In order to facilitate data analytics for big scholarly data, architectures and services for the same need to be developed. The evolving nature of research problems has made them essentially interdisciplinary. As a result, there is a growing demand for scholarly applications like collaborator discovery, expert finding and research recommendation systems, in addition to several others. This research paper investigates the current trends and identifies the existing challenges in development of a big scholarly data platform, with specific focus on directions for future research and maps them to the different phases of the big data lifecycle.

General information
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, Jamia Millia Islamia
Contributors: Khan, S., Liu, X., Shakil, K. A., Alam, M.
Pages: 923-944
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: Information Processing & Management
Volume: 53
Issue number: 4
ISSN (Print): 0306-4573
Ratings:
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 4.23 SJR 0.92 SNIP 2.666
Web of Science (2017): Impact factor 3.444
Web of Science (2017): Indexed yes
Original language: English
Keywords: Scholarly data, Big data, Cloud-based big data analysis, Big scholarly data, Big data platform, Scholarly applications
Electronic versions:
1606.01808..pdf. Embargo ended: 14/04/2019
DOIs:
10.1016/j.ipm.2017.03.006
Source: FindIt
Source ID: 2357007108
Research output: Contribution to journal › Journal article – Annual report year: 2017 › Research › peer-review

CITIESData: a smart city data management framework

Smart city data come from heterogeneous sources including various types of the Internet of Things such as traffic, weather, pollution, noise, and portable devices. They are characterized with diverse quality issues and with different types of sensitive information. This makes data processing and publishing challenging. In this paper, we propose a framework to streamline smart city data management, including data collection, cleansing, anonymization, and publishing. The paper classifies smart city data in sensitive, quasi-sensitive, and open/public levels and then suggests different strategies to process and publish the data within these categories. The paper evaluates the framework using a real-world smart city data set, and the results verify its effectiveness and efficiency. The framework can be a generic solution to manage smart city data.

General information
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, Department of Civil Engineering, Section for Building Energy, Centre for IT-Intelligent Energy Systems in Cities
Contributors: Liu, X., Heller, A., Nielsen, P. S.
Pages: 699-722
Publication date: 2017
Peer-reviewed: Yes
Collaborative Filtering Fusing Label Features Based on SDAE

Collaborative filtering (CF) is successfully applied to recommendation system by digging the latent features of users and items. However, conventional CF-based models usually suffer from the sparsity of rating matrices which would degrade model's recommendation performance. To address this sparsity problem, auxiliary information such as labels are utilized. Another approach of recommendation system is content-based model which can't be directly integrated with CF-based model due to its inherent characteristics. Considering that deep learning algorithms are capable of extracting deep latent features, this paper applies Stack Denoising Auto Encoder (SDAE) to content-based model and proposes LCF (Deep Learning for Collaborative Filtering) algorithm by combing CF-based model which fuses label features. Experiments on real-world data sets show that DLCF can largely overcome the sparsity problem and significantly improves the state of art approaches.

SciCloud: A Scientific Cloud and Management Platform for Smart City Data

The pervasive use of Internet of Things and smart meter technologies in smart cities increases the complexity of managing the data, due to their sizes, diversity, and privacy issues. This requires an innovate solution to process and manage the data effectively. This paper presents an elastic private scientific cloud, SciCloud, to tackle these grand challenges. SciCloud provides on-demand computing resource provisions, a scalable data management platform and an in-place data analytics environment to support the scientific research using smart city data.
A hybrid ICT-solution for smart meter data analytics

Smart meters are increasingly used worldwide. Smart meters are the advanced meters capable of measuring energy consumption at a fine-grained time interval, e.g., every 15 min. Smart meter data are typically bundled with social economic data in analytics, such as meter geographic locations, weather conditions and user information, which makes the data sets very sizable and the analytics complex. Data mining and emerging cloud computing technologies make collecting, processing, and analyzing the so-called big data possible. This paper proposes an innovative ICT-solution to streamline smart meter data analytics. The proposed solution offers an information integration pipeline for ingesting data from smart meters, a scalable platform for processing and mining big data sets, and a web portal for visualizing analytics results. The implemented system has a hybrid architecture of using Spark or Hive for big data processing, and using the machine learning toolkit, MADlib, for doing in-database data analytics in PostgreSQL database. This paper evaluates the key technologies of the proposed ICT-solution, and the results show the effectiveness and efficiency of using the system for both batch and online analytics.

A Prediction-based Smart Meter Data Generator

With the prevalence of cloud computing and Internet of Things (IoT), smart meters have become one of the main components of smart city strategy. Smart meters generate large amounts of fine-grained data that is used to provide useful information to consumers and utility companies for decision-making. Now-a-days, smart meter analytics systems consist of analytical algorithms that process massive amounts of data. These analytics algorithms require ample amounts of realistic data for testing and verification purposes. However, it is usually difficult to obtain adequate amounts of realistic data, mainly due to privacy issues. This paper proposes a smart meter data generator that can generate realistic energy consumption data by making use of a small real-world dataset as seed. The generator generates data using a prediction-based method that depends on historical energy consumption patterns along with Gaussian white noise. In this paper, we
comprehensively evaluate the efficiency and effectiveness of the proposed method based on a real-world energy data set.

**General information**
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, DTU Climate Centre, University College of Northern Denmark, Technical University of Denmark
Contributors: Iftikhar, N., Liu, X., Nordbjerg, F. E., Danalachi, S.
Number of pages: 8
Publication date: 2016

**Host publication information**
Title of host publication: Proceedings of the 19th International Conference on Network-Based Information Systems (NBiS), 2016
Publisher: IEEE
ISBN (Electronic): 978-1-5090-0979-4
DOI: 10.1109/NBiS.2016.15
Source: PublicationPreSubmission
Source ID: 125553676
Research output: Chapter in Book/Report/Conference proceeding – Article in proceedings – Annual report year: 2016 – Research – peer-review

**A Weighted K-AP Query Method for RSSI based Indoor Positioning**
The paper studies the establishment of offline fingerprint library based on RSSI (Received Signal Strength Indication), and proposes WF-SKL algorithm by introducing the correlation between RSSIs. The correlations can be transformed as AP fingerprint sequence to build the offline fingerprint library. To eliminate the positioning error caused by instable RSSI value, WF-SKL can filter the noise AP via online AP selection, meanwhile it also reduces the computation load. WF-SKL utilizes LCS algorithm to find out the measurement between the nearest neighbors, and it proposes K-AP (P,Q) nearest neighbor queries between two sets based on Map-Reduce framework. The algorithm can find out K (P,Q) nearest positions and weighted them for re-positioning to accelerate the matching speed between online data and offline data, and also improve the efficiency of positioning. According to a large scale positioning experiments, WF-SKL algorithm proves its high accuracy and positioning speed comparing with KNN indoor positioning.

**CITIESData: Towards Cloud Based Big Data Management for Smart Cities**
**General information**
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, DTU Climate Centre, Department of Civil Engineering, Section for Building Energy, Centre for IT-Intelligent Energy Systems in Cities
Contributors: Liu, X., Heller, A., Nielsen, P. S.
Number of pages: 1
Publication date: 2016
Peer-reviewed: No
Event: Poster session presented at 3rd General Consortium Meeting of Smart Cities project, CITIES, Kgs. Lyngby, Denmark.
Online Anomaly Energy Consumption Detection Using Lambda Architecture

With the widely use of smart meters in the energy sector, anomaly detection becomes a crucial mean to study the unusual consumption behaviors of customers, and to discover unexpected events of using energy promptly. Detecting consumption anomalies is, essentially, a real-time big data analytics problem, which does data mining on a large amount of parallel data streams from smart meters. In this paper, we propose a supervised learning and statistical-based anomaly detection method, and implement a Lambda system using the in-memory distributed computing framework, Spark and its extension Spark Streaming. The system supports not only iterative refreshing the detection models from scalable data sets, but also real-time anomaly detection on scalable live data streams. This paper empirically evaluates the system and the detection algorithm, and the results show the effectiveness and the scalability of the lambda detection system.

Optimizing ETL by a Two-level Data Staging Method

In data warehousing, the data from source systems are populated into a central data warehouse (DW) through extraction, transformation and loading (ETL). The standard ETL approach usually uses sequential jobs to process the data with dependencies, such as dimension and fact data. It is a non-trivial task to process the so-called early-/late-arriving data, which arrive out of order. This paper proposes a two-level data staging area method to optimize ETL. The proposed method is an all-in-one solution that supports processing different types of data from operational systems, including early-/late-arriving data, and fast-/slowly-changing data. The introduced additional staging area decouples loading process from data extraction and transformation, which improves ETL flexibility and minimizes intervention to the data warehouse. This paper evaluates the proposed method empirically, which shows that it is more efficient and less intrusive than the standard ETL method.
Smart Meter Data Analytics: Systems, Algorithms and Benchmarking

Smart electricity meters have been replacing conventional meters worldwide, enabling automated collection of fine-grained (e.g., every 15 minutes or hourly) consumption data. A variety of smart meter analytics algorithms and applications have been proposed, mainly in the smart grid literature. However, the focus has been on what can be done with the data rather than how to do it efficiently. In this paper, we examine smart meter analytics from a software performance perspective. First, we design a performance benchmark that includes common smart meter analytics tasks. These include off-line feature extraction and model building as well a framework for on-line anomaly detection that we propose. Second, since obtaining real smart meter data is difficult due to privacy issues, we present an algorithm for generating large realistic data sets from a small seed of real data. Third, we implement the proposed benchmark using five representative platforms: a traditional numeric computing platform (Matlab), a relational DBMS with a built-in machine learning toolkit (PostgreSQL/MADlib), a main-memory column store (“System C”), and two distributed data processing platforms (Hive and Spark/Spark Streaming). We compare the five platforms in terms of application development effort and performance on a multicore machine as well as a cluster of 16 commodity servers.

General information
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, DTU Climate Centre, University of Waterloo
Contributors: Liu, X., Golab, L., Golab, W., Ilyas, I. F., Jin, S.
Number of pages: 38
Publication date: 2016
Peer-reviewed: Yes
Streamlining Meter Data Analytics

**General information**
Publication status: Published
Organisations: Department of Management Engineering, Systems Analysis, DTU Climate Centre, Centre for IT-Intelligent Energy Systems in Cities
Contributors: Liu, X., Nielsen, P. S.
Number of pages: 1
Publication date: 2016
Peer-reviewed: No
Event: Poster session presented at 3rd General Consortium Meeting of Smart Cities project, CITIES, Kgs. Lyngby, Denmark.
Electronic versions: smas_poster.pdf
URLs:
Research output: Contribution to conference › Poster – Annual report year: 2016 › Research

**An ETL Optimization Framework Using Partitioning and Parallelization**
Extract-Transform-Load (ETL) handles large amounts of data and manages workload through dataflows. ETL dataflows are widely regarded as complex and expensive operations in terms of time and system resources. In order to minimize the time and the resources required by ETL dataflows, this paper presents an optimization framework using partitioning and parallelization. The framework first partitions an ETL dataflow into multiple execution trees according to the characteristics of ETL constructs, then within an execution tree pipelined parallelism and shared cache are used to optimize the partitioned dataflow. Furthermore, multi-threading is used in component-based optimization. The experimental results show that the proposed framework can achieve 4.7 times faster than the ordinary ETL dataflows (without using the proposed partitioning and optimization methods), and is comparable to the similar ETL tools.

**General information**
Publication status: Published
Organisations: University of Waterloo, Aalborg University
Contributors: Liu, X., Iftikhar, N.
Number of pages: 8
Publication date: 2015

**Host publication Information**
Title of host publication: Proceedings of the 30th ACM Symposium on Applied Computing (SAC 2015)
Publisher: Association for Computing Machinery
ISBN (Print): 978-1-4503-3196-8
DOIs: 10.1145/2695664.2695846
Source: PublicationPreSubmission
Source ID: 110379354
Research output: Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report year: 2015 › Research › peer-review

**Benchmarking Smart Meter Data Analytics**
Smart electricity meters have been replacing conventional meters worldwide, enabling automated collection of fine-grained (every 15 minutes or hourly) consumption data. A variety of smart meter analytics algorithms and applications have been proposed, mainly in the smart grid literature, but the focus thus far has been on what can be done with the data rather than how to do it efficiently. In this paper, we examine smart meter analytics from a software performance perspective. First, we propose a performance benchmark that includes common data analysis tasks on smart meter data. Second, since obtaining large amounts of smart meter data is difficult due to privacy issues, we present an algorithm for generating large realistic data sets from a small seed of real data. Third, we implement the proposed benchmark using five representative platforms: a traditional numeric computing platform (Matlab), a relational DBMS with a built-in machine learning toolkit (PostgreSQL/ MADLib), a main-memory column store (“System C”), and two distributed data processing platforms (Hive and Spark). We compare the five platforms in terms of application development effort and performance on a multi-core machine as well as a cluster of 16 commodity servers. We have made the proposed benchmark and data generator freely available online.

**General information**
Publication status: Published
Organisations: University of Waterloo
Contributors: Liu, X., Golab, L., Golab, W., Ilyas, I. F.
Cloud-Based Software Platform for Smart Meter Data Management

Today smart meters are increasingly used in the worldwide. Smart meters are the advanced meters capable of measuring customer energy consumption at a fine-grained time interval, e.g., every 15 minutes. The data are very sizeable, and might be from different sources, along with the other social-economic metrics such as the geographic information of meters, the information about users and their property, geographic location and others, which make the data management very complex. On the other hand, data-mining and the emerging cloud computing technologies make the collection, management, and analysis of the so-called big data possible. This can improve energy management, e.g., help utility companies to forecast energy loads and improve services, and help households to manage energy usage and save money. As this regard, the proposed paper focuses on building an innovative software platform for smart meter data analytics using cloud technologies, aiming to maximize the information assets in demand-side energy management and relieving peak load. The proposed platform will offer information integration pipeline to ingest smart meter time-series data; a secure repository for researchers sharing their knowledge; scalable data analytics platform for data mining over big data sets for energy demand forecasting and consumption discovering; data as the service for other applications using smart meter data; and a portal for visualizing data analytics results. The design will incorporate hybrid clouds, including Infrastructure as a Service (IaaS) and Platform as a Service (PaaS), which are suitable for on-demand provisioning, massive scaling, and manageability. Besides, the design will impose extensibility, efficiency, and high availability on the system. The paper will evaluate the system comprehensively, and compare with other similar works. This paper will provide a proof of concept for building the data management system expanding from the data management of energy sector to the entire sectors of smart cities.

Relational-Based Sensor Data Cleansing

Today sensors are widely used in many monitoring applications. Due to some random environmental effects and/or sensing failures, the collected sensor data is typically noisy. Thus, it is critical to cleanse the sensor data before using it to answer queries or conduct data analysis. Popular data cleansing approaches, such as classification, prediction and moving average are not suited for embedded sensor devices, due to the limited storage and processing capabilities. In this paper, we propose a sensor data cleansing approach using the relational-based technologies, including constraints, triggers and granularity-based data aggregation. The proposed approach is simple but effective to cleanse different types of dirty data, including delayed data, incomplete data, incorrect data, duplicate data and missing data. We evaluate the
proposed strategy to verify its efficiency, effectiveness and adaptability.

**General information**
Publication status: Published  
Organisations: Department of Management Engineering, Systems Analysis, DTU Climate Centre, Energy Systems Analysis, University College of Northern Denmark  
Contributors: Iftikhar, N., Liu, X., Nordbjerg, F. E.  
Number of pages: 14  
Publication date: 2015

**Host publication information**
Title of host publication: Proceedings. 19th East-European Conference on Advances in Databases and Information Systems  
Keywords: Data cleansing, Sensor data, Relational-based, Simple, Effective  
Source: PublicationPreSubmission  
Source ID: 110808682  
Research output: Chapter in Book/Report/Conference proceeding › Article in proceedings – Annual report year: 2015 › Research › peer-review

SMAS: Smart Meter Data Analytics System
Smart electricity meters are replacing conventional meters worldwide and have enabled a new application domain: smart meter data analytics. In this paper, we introduce SMAS, our smart meter analytics system, which demonstrates the actionable insight that consumers and utilities can obtain from smart meter data. Notably, we implemented SMAS inside a relational database management system using open source tools: PostgreSQL and the MADLib machine learning toolkit. In the proposed demonstration, conference attendees will interact with SMAS as electricity providers, consultants and consumers, and will perform various analyses on real data sets.

**General information**
Publication status: Published  
Organisations: University of Waterloo  
Contributors: Liu, X., Golab, L., Ilyas, I. F.  
Number of pages: 4  
Publication date: 2015  
Peer-reviewed: Yes  
Electronic versions:  
SMAS.pdf  
Source: PublicationPreSubmission  
Source ID: 110339890  
Research output: Contribution to conference › Paper – Annual report year: 2015 › Research › peer-review

Streamlining Smart Meter Data Analytics
Today smart meters are increasingly used in worldwide. Smart meters are the advanced meters capable of measuring customer energy consumption at a fine-grained time interval, e.g., every 15 minutes. The data are very sizable, and might be from different sources, along with the other social-economic metrics such as the geographic information of meters, the information about users and their property, geographic location and others, which make the data management very complex. On the other hand, data-mining and the emerging cloud computing technologies make the collection, management, and analysis of the so-called big data possible. This can improve energy management, e.g., help utilities improve the management of energy and services, and help customers save money. As this regard, the paper focuses on building an innovative software solution to streamline smart meter data analytic, aiming at dealing with the complexity of data processing and data analytics. The system offers an information integration pipeline to ingest smart meter data; scalable data processing and analytic platform for pre-processing and mining big smart meter data sets; and a web-based portal for visualizing data analytics results. The system incorporates hybrid technologies, including big data technologies Spark and Hive, the high performance RDBMS PostgreSQL with the in-database machine learning toolkit, MADlib, which are able to satisfy a variety of requirements in smart meter data analytics.

**General information**
Publication status: Published  
Organisations: Department of Management Engineering, Systems Analysis, DTU Climate Centre, Energy Systems Analysis, Centre for IT-Intelligent Energy Systems in Cities  
Contributors: Liu, X., Nielsen, P. S.  
Number of pages: 14  
Publication date: 2015

**Host publication information**
CloudETL: Scalable Dimensional ETL for Hive

Extract-Transform-Load (ETL) programs process data into data warehouses (DWs). Rapidly growing data volumes demand systems that scale out. Recently, much attention has been given to MapReduce for parallel handling of massive data sets in cloud environments. Hive is the most widely used RDBMS-like system for DWs on MapReduce and provides scalable analytics. It is, however, challenging to do proper dimensional ETL processing with Hive; e.g., the concept of slowly changing dimensions (SCDs) is not supported (and due to lacking support for UPDATEs, SCDs are complex to handle manually). Also the powerful Pig platform for data processing on MapReduce does not support such dimensional ETL processing. To remedy this, we present the ETL framework CloudETL which uses Hadoop to parallelize ETL execution and to process data into Hive. The user defines the ETL process by means of high-level constructs and transformations and does not have to worry about technical MapReduce details. CloudETL supports different dimensional concepts such as star schemas and SCDs. We present how CloudETL works and uses different performance optimizations including a purpose-specific data placement policy to co-locate data. Further, we present a performance study and compare with other cloud-enabled systems. The results show that CloudETL scales very well and outperforms the dimensional ETL capabilities of Hive both with respect to performance and programmer productivity. For example, Hive uses 3.9 times as long to load an SCD in an experiment and needs 112 statements while CloudETL only needs 4.

Survey of Real-time Processing Systems for Big Data

In recent years, real-time processing and analytics systems for big data—in the context of Business Intelligence (BI)—have received a growing attention. The traditional BI platforms that perform regular updates on daily, weekly or monthly basis are no longer adequate to satisfy the fast-changing business environments. However, due to the nature of big data, it has become a challenge to achieve the real-time capability using the traditional technologies. The recent distributed computing technology, MapReduce, provides off-the-shelf high scalability that can significantly shorten the processing time for big data. Its open-source implementation such as Hadoop has become the de-facto standard for processing big data, however, Hadoop has the limitation of supporting real-time updates. The improvements in Hadoop for the real-time capability, and the other alternative real-time frameworks have been emerging in recent years. This paper presents a survey of the open source technologies that support big data processing in a real-time/near real-time fashion, including their system architectures and platforms.
ETLMR: A Highly Scalable Dimensional ETL Framework Based on MapReduce

Extract-Transform-Load (ETL) flows periodically populate data warehouses (DWs) with data from different source systems. An increasing challenge for ETL flows is to process huge volumes of data quickly. MapReduce is establishing itself as the de-facto standard for large-scale data-intensive processing. However, MapReduce lacks support for high-level ETL specific constructs, resulting in low ETL programmer productivity. This paper presents a scalable dimensional ETL framework, ETLMR, based on MapReduce. ETLMR has built-in native support for operations on DW-specific constructs such as star schemas, snowflake schemas and slowly changing dimensions (SCDs). This enables ETL developers to construct scalable MapReduce-based ETL flows with very few code lines. To achieve good performance and load balancing, a number of dimension and fact processing schemes are presented, including techniques for efficiently processing different types of dimensions. The paper describes the integration of ETLMR with a MapReduce framework and evaluates its performance on large realistic data sets. The experimental results show that ETLMR achieves very good scalability and compares favourably with other MapReduce data warehousing tools.

Ontology-Based Big Dimension Modeling in Data Warehouse Schema Design

During data warehouse schema design, designers often encounter how to model big dimensions that typically contain a large number of attributes and records. To investigate effective approaches for modeling big dimensions is necessary in order to achieve better query performance, with respect to response time. In most cases, the big dimension modeling process is complicated since it usually requires accurate description of business semantics, multiple design revisions and comprehensive testings. In this paper, we present the design methods for modeling big dimensions, which include horizontal partitioning, vertical partitioning and their hybrid. We formalize the design methods, and propose an algorithm that describes the modeling process from an OWL ontology to a data warehouse schema. In addition, this paper also presents an effective ontology-based tool to automate the modeling process. The tool can automatically generate the data warehouse schema from the ontology of describing the terms and business semantics for the big dimension. In case of any change in the requirements, we only need to modify the ontology, and re-generate the schema using the tool. This paper also evaluates the proposed methods based on sample sales data mart.
3XL: An Efficient DBMS-Based Triple-Store

This paper demonstrates the use of 3XL, a DBMS-based triple-store for OWL Lite data. 3XL is characterized by its use of a database schema specialized for the data to represent. The specialized database schema uses object-relational features — particularly inheritance — and partitions the data such that it is fast to locate the needed data when it is queried. Further, the generated database schema is very intuitive and it is thus easy to integrate the OWL data with other kinds of data. 3XL offers performance comparable to the leading file-based triple-stores. We will demonstrate 1) how a specialized database schema is generated by 3XL based on an OWL ontology; 2) how triples are loaded, including how they pass through the 3XL system and how 3XL can be configured to fine-tune performance; and 3) how (simple and complex) queries can be expressed and how they are executed by 3XL.

Data warehousing technologies for large-scale and right-time data

This thesis is about data warehousing technologies for large-scale and right-time data. Today, due to the exponential growth of data, it has become a common practice for many enterprises to process hundreds of gigabytes of data per day. Traditionally, data warehousing populates data from heterogeneous sources into a central data warehouse (DW) by Extract-Transform-Load (ETL) at regular time intervals, e.g., monthly, weekly, or daily. But now, it becomes challenging for large-scale data, and hard to meet the near real-time/right-time business decisions. This thesis considers some of these challenges and makes the following contributions: First, this thesis presents a new and efficient way to store triples from an OWL Lite ontology known from the Semantic Web field. In contrast to classic triple-stores where the data with the triple format of (subject; predicate; object) is stored in few, but big, tables with few columns, the presented triple-store spreads the data over more tables that may have many columns. The triple-store is optimized by an extensive use of bulk techniques, which makes it very efficient to insert and extract data. The DBMS-based solution makes it very flexible to integrate with other non-triple data. Second, this thesis presents a middleware system for live DW data. Processing live DW data is one of the most tricky problems in data warehousing. An innovative method is proposed for processing live DW data, which accumulates the data in an intermediate data store, and does data modifications on-the-fly when the data is materialized or queried. The data is made available in the DW exactly when needed and users can get bulk-load speeds, but INSERT-like data availability. Third, this thesis presents the first dimensional ETL programming framework using MapReduce. Parallel ETL is needed for large-scale data, but it is not easy to implement. This presented framework makes this very easy by offering high-level ETL-specific constructs, including those for star schema, snowflake schema, slowly changing dimensions (SCDs) and very large dimensions. The framework can achieve high programming efficiency, i.e., only a few statements needed for implementing a parallel ETL program, and good scalability for processing different DW schemas. Finally, this thesis presents scalable dimensional ETL for cloud warehouse. Today, organizations gain growing interest in moving data warehousing systems towards the cloud, however, the current data warehousing systems are not yet particularly suited for the cloud. The presented framework exploits Hadoop to parallelize ETL execution and Hive as the warehouse system. It has a shared-nothing architecture, and supports scalable ETL operations on clustered commodity machines. To implement dimensional ETL, this framework can achieve higher programmer productivity than Hive, and its performance is also better. In summary, this thesis discusses several aspects of the current challenges and problems of data warehousing, including integrating Web data, near real-time/right-time data warehousing, handling the exponential growth of data and cloud data warehousing. This thesis proposes a variety of technologies to deal with these specific issues.
MapReduce-based Dimensional ETL Made Easy
This paper demonstrates ETLMR, a novel dimensional Extract–Transform–Load (ETL) programming framework that uses MapReduce to achieve scalability. ETLMR has built-in native support of data warehouse (DW) specific constructs such as star schemas, snowflake schemas, and slowly changing dimensions (SCDs). This makes it possible to build MapReduce-based dimensional ETL flows very easily. The ETL process can be configured with only few lines of code. We will demonstrate the concrete steps in using ETLMR to load data into a (partly snowflaked) DW schema. This includes configuration of data sources and targets, dimension processing schemes, fact processing, and deployment. In addition, we also present the scalability on large data sets.

3XL: Supporting efficient operations on very large OWL Lite triple-stores
An increasing number of (semantic) web applications store a very large number of (subject, predicate, object) triples in specialized storage engines called triple-stores. Often, triple-stores are used mainly as plain data stores, i.e., for inserting and retrieving large amounts of triples, but not using more advanced features such as logical inference, etc. However, current triple-stores are not optimized for such bulk operations and/or do not support OWL Lite. Further, triple-stores can be inflexible when the data has to be integrated with other kinds of data in non-triple form, e.g., standard relational data. This paper presents 3XL, a triple-store that efficiently supports operations on very large amounts of OWL Lite triples. 3XL also provides the user with high flexibility as it stores data in an object-relational database in a schema that is easy to use and understand. It is, thus, easy to integrate 3XL data with data from other sources. The distinguishing features of 3XL include (a) flexibility as the data is stored in a database, allowing easy integration with other data, and can be queried by means of both triple queries and SQL, (b) using a specialized data-dependent schema (with intelligent partitioning) which is intuitive and efficient to use, (c) using object-relational DBMS features such as inheritance, (d) efficient loading through extensive use of bulk loading and caching, and (e) efficient triple query operations, especially in the important case when the subject and/or predicate is known. Extensive experiments with a PostgreSQL-based implementation show that 3XL performs very well for such operations and that the performance is comparable to state-of-the-art triple-stores.
ETLMR: A Highly Scalable Dimensional ETL Framework Based on MapReduce

Extract-Transform-Load (ETL) flows periodically populate data warehouses (DWs) with data from different source systems. An increasing challenge for ETL flows is processing huge volumes of data quickly. MapReduce is establishing itself as the de-facto standard for large-scale data-intensive processing. However, MapReduce lacks support for high-level ETL specific constructs, resulting in low ETL programmer productivity. This paper presents a scalable dimensional ETL framework, ETLMR, based on MapReduce. ETLMR has built-in native support for operations on DW-specific constructs such as star schemas, snowflake schemas and slowly changing dimensions (SCDs). This enables ETL developers to construct scalable MapReduce-based ETL flows with very few code lines. To achieve good performance and load balancing, a number of dimension and fact processing schemes are presented, including techniques for efficiently processing different types of dimensions. The paper describes the integration of ETLMR with a MapReduce framework and evaluates its performance on large realistic data sets. The experimental results show that ETLMR achieves very good scalability and compares favourably with other MapReduce data warehousing tools.
A Data Warehouse Solution for e-Government

The eGovMon Data Warehouse (eGovMon DW) is built as a data repository for eGovernment services benchmarking results. We propose a DW architecture with open source business intelligence technologies for eGovernment. This DW architecture uses PostgreSQL as the DBMS, eGovernment operational system as the data source, and a right-time ETL tool to populate the data. Through this proposal, we give the potential research interests and issues for our future work.

Online Testing of the ABOT Game Server Using the Qtronic Tool

This technical report describes our experience in online testing using Conformiq Qtronic, a model-based testing tool that can automate the test generation and execution. A mobile game server is used as system under test (SUT). We show how the server is modeled in Qtronic, how an adapter between Qtronic and the SUT is implemented, and how the automatically generated tests are run against the SUT. The approach proved that taking advantage of the automatically generated tests, the SUT can be tested extensively within a short time frame, thus increasing the chances of discovering hidden implementation errors.
Flexible energy Denmark
Liu, X., PI, Department of Technology, Management and Economics, Sustainability, Energy Systems Analysis
01/04/2019 → 17/09/2019
Project: Research

FlexSUS : FlexSUS - Flexibility for Smart Urban Energy Systems
The FlexSUS project develops a decision-making tool to support smart urban energy systems based on
digital solutions which enable municipalities and city planners to optimize their energy systems while
implementing climate change mitigation efforts. The project builds on optimisation of existing and planned
energy systems and develops bridges across them in the form of data management system and
visualization tool to be used by system planners that adapt to the specific needs and energy resources of
the cities.
Bergaentzlé, C., Project Coordinator, Energy Economics and Regulation, Sustainability, Department of Technology,
Management and Economics
Papakonstantinou, A., Project Manager, Energy Economics and Regulation, Sustainability, Department of Technology,
Management and Economics
Nielsen, P. S., Project Participant, Sustainability, Energy Systems Analysis, Department of Technology, Management and
Economics
Liu, X., Project Participant, Sustainability, Energy Systems Analysis, Department of Technology, Management and
Economics
Bolwig, S., Project Participant, Innovation, Industrial Dynamics, Department of Technology, Management and Economics
Gregg, J. S., Project Participant, Innovation, Industrial Dynamics, Department of Technology, Management and Economics
McKenna, R., Project Participant, Sustainability, Energy Systems Analysis, Department of Technology, Management and
Economics
Project: Research

Data Intelligence Approaches for Energy Peak Management
Dai, W., PhD Student, Department of Management Engineering
Liu, X., Main Supervisor
Nielsen, P. S., Supervisor
15/04/2019 → 14/04/2022
Project: PhD

Roskilde Smart Monitoring Household project
Monitoring fine-grained residential household energy consumption and resident activities
Liu, X., PI, Department of Management Engineering, Systems Analysis
01/02/2018 → 30/09/2019
Project: Research

SciCloud: Scientific Cloud
Liu, X., Project Participant, Department of Management Engineering, Systems Analysis
01/01/1970 → 30/06/2018
Project: Research

Roskilde Smart Monitoring Household
Liu, X., Project Participant, Department of Management Engineering, Systems Analysis
01/01/2018 → …
Project: Research

CTT: Carbon Track and Trace
Liu, X., Project Participant, Department of Management Engineering, Systems Analysis
10/09/2016 → 30/09/2018
Project: Research

SCA: Smart City Accelerator
Smart Cities Accelerator (SCA) is financed by the EU programme Interreg-ØKS, that supports crossregional collaboration
in Europe. The project runs for a period of three years 2016 – 2019, and is supported by 6.468.000 Euro, where 50 % is
cofinanced by the 11 partners. Climate challenges such as global warming, air pollution and pressure on natural resources has resulted in a number of political objectives at local/regional, national, European and global level. One of the major climate objectives is to substitute fossil fuels with sustainable and renewable energy. This requires interdisciplinary development of methods and technologies. SCA focuses on facilitating knowledge sharing and the development of demonstration projects that can create more sustainable solutions within the municipal energy supply system from energy production to energy consumption.

SCA will facilitate the development of greener solutions by combining seven vertical focus areas within the energy sector with four cross sectorial horizontal areas. The horizontal areas are anchored in strong research environments including data, behaviour, law/regulations and learnings on both sides of the Sound.

Gregg, J. S., Project Manager, Department of Management Engineering, Systems Analysis Nielsen, P. S., Project Participant, Department of Management Engineering, Systems Analysis Liu, X., Project Participant, Department of Management Engineering, Systems Analysis Tanner, A. N., Project Participant, Department of Management Engineering, Technology and Innovation Management Rosati, F., Project Participant, Department of Management Engineering, Technology and Innovation Management
01/07/2017 → 01/06/2020
Keywords: smart cities, sustainability assessment, social cohesion
Project: Research

**ClairCity: ClairCity**

ClairCity is an innovative project involving thousands of people in cities across Europe, enabling us all to decide the best local options for a future with clean air and lower carbon emissions. ClairCity is funded by the European Union.

Kewo, A., Project Participant, Department of Management Engineering, Systems Analysis Nielsen, P. S., Project Participant, Department of Management Engineering, Systems Analysis Liu, X., Project Participant, Department of Management Engineering, Systems Analysis
01/07/2017 → 31/01/2019
Keywords: Temporal_resolution, Spatiotemporal_resolution, Energy modelling
Project: Research

**Benchmarking Residential Energy Consumption In Indonesia**

Kewo, A., PhD Student, Department of Management Engineering Nielsen, P. S., Main Supervisor Liu, X., Supervisor
Stipendie fra udlandet
01/05/2016 → 30/04/2020
Award relations: Benchmarking Residential Energy Consumption In Indonesia
Project: PhD

**CITIES**

Liu, X., Project Participant, Department of Management Engineering, Systems Analysis, DTU Climate Centre, Energy Systems Analysis
01/02/2015 → 31/01/2017
Project: Research

**Activities:**

**User Activity Simulation for Residential Buildings**

Period: 2 Oct 2018
Xiufeng Liu (Guest lecturer)
Department of Management Engineering Systems Analysis

Description
conference presentation

Related event

**13th CONFERENCE ON SUSTAINABLE DEVELOPMENT OF ENERGY WATER AND ENVIRONMENT SYSTEMS**
01/10/2018 → 04/10/2018
Palermo, Italy
The Architecture Development of Smart Meter Controlling System in dynamic IP addresses environment  
Period: 1 Mar 2018  
Angreine Kewo (Speaker)  
Per Sieverts Nielsen (Other)  
Xiufeng Liu (Other)  
Department of Management Engineering  
Systems Analysis  
Degree of recognition: International  
Links:  
http://International Conference on Intelligent Autonomous Systems

Related event  
International Conference on Intelligent Autonomous Systems  
01/03/2018 → 03/03/2018  
Singapore, Singapore  
Keywords: Control, autonomous  
Activity: Talks and presentations › Conference presentations

Analysis and Visualization of Urban Emission Measurements in Smart Cities  
Period: Mar 2018  
Xiufeng Liu (Guest lecturer)  
Department of Management Engineering  
Systems Analysis  
Description  
EDBT conference  
Related event  
21st International Conference on Extending Database Technology (EDBT)  
26/03/2018 → 29/03/2018  
Vienna, Austria  
Activity: Talks and presentations › Conference presentations

Modelling of electricity consumption in one of the world’s most populous cities–Jakarta, Indonesia  
Period: 4 Oct 2017  
Angreine Kewo (Speaker)  
Per Sieverts Nielsen (Other)  
Xiufeng Liu (Other)  
Department of Management Engineering  
Systems Analysis  
Degree of recognition: International  
Related event  
12th sdawes Conference  
04/10/2017 → 08/10/2017  
Dubrovnik, Croatia  
Activity: Talks and presentations › Conference presentations

18th International Conference on Big Data Analytics and Knowledge Discovery (DaWaK 2016)  
Period: 6 Sep 2016  
Xiufeng Liu (Participant)
Department of Management Engineering
Systems Analysis
DTU Climate Centre

Related event

18th International Conference on Big Data Analytics and Knowledge Discovery (DaWaK 2016)
05/09/2016 → 08/09/2016
Porto, Portugal
Activity: Attending an event › Participating in or organising a conference

10th Conference on Sustainable Development of Energy, Water and Environment Systems
Period: 1 Feb 2015
Xiufeng Liu (Participant)
Department of Management Engineering
Systems Analysis
DTU Climate Centre
Energy Systems Analysis
Centre for IT-Intelligent Energy Systems in Cities

Related event

10th Conference on Sustainable Development of Energy, Water and Environment Systems
27/09/2015 → 02/10/2015
Dubrovnik, Croatia
Activity: Attending an event › Participating in or organising a conference