

2014 International Conference on Data Science & Engineering

(ICDSE 2014)

**Kochi, India
26-28 August 2014**



**IEEE Catalog Number: CFP1427T-POD
ISBN: 978-1-4799-5462-9**

Table of Contents

ICDSE 2014

2014 International Conference on Data Science & Engineering

Welcome Message

Conference Organization

Session 1: Data Processing on Big Data Infrastructure

Efficient Structuring of Data in Big Data	1
<i>Ashwin Kumar T K, Hong Liu and Johnson P Thomas</i>	
Data Driven Mobile Commerce Intelligence with Real-time Advanced Inexpensive Network Computing	6
<i>K. Rajasekhar and Niraj Upadhayaya</i>	
Dynamic Incremental Maintenance of Materialized View Based on Attribute Affinity	12
<i>Partha Ghosh and Soumya Sen</i>	
SQL-MapReduce Hybrid Approach towards Distributed Projected Clustering	18
<i>Sandhya Harikumar, M. Shyju and M.R. Kaimal</i>	
Fault Tolerant State Management for High-Volume Low-Latency Data Stream Workloads	24
<i>K.B.Muralidharan, G.Santhosh Kumar and M.Bhasi</i>	
An Impudent Approach for Prudential Hadoop Cluster	28
<i>Anjali P P, Bincy P Andrews, Sreelakshmi Ganesh, Suman Elizabeth Daniel and Binu A</i>	
Statistical Analysis to Determine the Performance of Multiple Beneficiaries of Educational Sector Using Hadoop-Hive	32
<i>Sreelakshmi Ganesh and Binu A</i>	

Session 2: Cloud Computing and Data Management

- Advancing Data Management and Usage in a Major Australian Health Service: The REASON Discovery Platform **38**

Christopher Bain and Chris Mac Manus

- Cloud Light Weight: a New Solution for Load Balancing in Cloud Computing **44**

Mohammadreza Mesbahi, Amir Masoud Rahmani and

Anthony Theodore Chronopoulos

- Game Theoretic Approach to Resource Provisioning in a Distributed Cloud **51**

Khethavath Praveen, Johnson Thomas and Hong Liu

- Analysis of Job Execution Reliability in a Grid through Job Accounting Tool **57**

Jibin Sunny, Divya MG and Subrata Chattopadhyay

- A Robust Ticket-Based Mutual Authentication Scheme for Data Security in Cloud Computing **62**

Mrudula Sarvabhatla, M.Giri and Chandra Sekhar Vorugunti

- Knowledge Driven Decision Support System for Provenance Models in Relational Database **68**

Asma Rani and Sunil Thalia

Session 3: Machine Learning Techniques and Systems

- A Nature-inspired Hybrid Fuzzy C-means Algorithm for Better Clustering of Biological Data Sets **76**

Arunanand T A, Abdul Nazeer K A, Mathew J Palakal and

Meeta Pradhan

- DeMalFier: Detection of Malicious Web Pages using an Effective Classifier **83**

Asha S Manek, Sumithra V, P Deepa Shenoy, M. Chandra Mohan, Venugopal K R and L M Patnaik

An Improved ICPACA Based K-means Algorithm with Self Determined Centroids	89
<i>Christina Jacob and K A Abdul Nazeer</i>	
Sustainable Concrete Mix-Design: Evolutionary Algorithm Approach	94
<i>Saikat Dan and S.V. Barai</i>	
Improved Document Classification Through Enhanced Naive Bayes Algorithm	100
<i>Shiju Sathyadevan, Athira U, Sarath P R and Anjana V</i>	
Semantic Integration of Heterogeneous Relational Schemas using Multiple L1 Linear Regression and SVD	105
<i>Sandhya Harikumar, R. Reethima and M.R. Kaimal</i>	
Fuzzy Inference System based on Entropy of Traffic or Bot Detection on an Endpoint Host	112
<i>Soniya B and M Wilscy</i>	
A Concept based Clustering Model for Document Similarity	118
<i>Veena G and Lekha N K</i>	
Session 4: Information Storage, Search and Retrieval	
CancerEpiDB: A Database for DNA methylated Genes in Cancer	124
<i>Akshay Desai, Mathew Palakal, Meeta Pradhan and Premkumar Duraiswamy</i>	
Adaptive Hashing Based Multiple Variable Length Pattern Search Algorithm for Large Data Sets	130
<i>Punit Kanuga and Anamika Chauhan</i>	
Incremental Multi-Document Summarization: An Incremental Clustering Based Approach	136
<i>Johne John and Asharaf S</i>	

A Time Saving Index Construction in Geospatial Searching	140
<i>Amruta Joshi and Ujjwala Patil</i>	

Text Summarization for Malayalam Documents - an Experience	145
<i>Rajina Kabeer and Sumam Mary Idicula</i>	

Session 5: Social Networks and Personal Information

Affiliations network analysis in scientific citations: A Case study of Information Technology for Engineering	151
<i>Hiran H. Lathabai, Thara Prabhakaran and Manoj Changat</i>	

Effective Personalized Mobile Search using KNN	157
<i>K. Swati and A. J. Patankar</i>	

Modeling Diffusion Pattern in Online Social Media	161
<i>Mini U and Poulose Jacob. K</i>	

Session 6: Modeling and Simulation

Modeling of Rice Productivity in India	166
<i>Vidya Vinodini M.D, Reshmi Nair L.R, Aiswaria Lakshmi K.G,</i>	
<i>Anitha V.R, Reshma P, R Ajith Kumar, T.K Manoj kumar</i>	

Multivariate Regression Models for Prediction of Wind Speed	171
<i>Arjun N.N, Prema V., Krishna Kumar D., Prashanth P.,</i>	
<i>Sumantha Preeekshit V. and Uma Rao K</i>	

Multivariate Regression for Prediction of Solar Irradiance	177
<i>Nalina U., Prema V, Smitha K, and Uma Rao K</i>	

Implementation of a Novel SCADA Architecture for a 210 MW Thermal Power Plant	182
<i>Jerry Daniel J, Pratheesh H, Raja Singh B, Rakesh G,</i>	
<i>Joseph Mathew, Senju Thomas Panicker, Shankar S S, Sujimon K T,</i>	
<i>Arun Krishnan and Muhammed Irshad A</i>	

Session 7: Computer Networks, Privacy and Security

A Communication Architecture Using Low Frequency Signals	188
<i>Kuruville Mathew, Biju Issac and Tan Chong Eng</i>	
Traffic Capture beyond 10 Gbps: Linear Scaling with Multiple Network Interface Cards on Commodity Servers	194
<i>Vinu Paul M V, Raktim Bhattacharjee and Rajesh R</i>	
Disclosure Risk of Individuals: A k-anonymity Study on Health Care Data Related to Indian Population	200
<i>Jisha Jose Panackal, Anitha S Pillai and V N Krishnachandran</i>	
A Modified Remote User Authentication Scheme using Smart Cards with Check Digits	206
<i>Shipra Kumari and Hari Om</i>	
An Effective Private Data Storage and Retrieval System using Secret Sharing Scheme based on Secure Multi-party Computation	210
<i>Divya G Nair, Binu.V.P and G.Santhosh Kumar</i>	

Session 8: Image Processing and Computer Vision

Image Denoising with True Pixel Presence Detection Procedure	215
<i>Sruthi M, Mredhula L and Dorairangaswamy M A</i>	
Recognition of Hand Gestures of English Alphabets using HOG Method	219
<i>Daleesha M Viswanathan and Sumam Mary Idicula</i>	
Alzheimer's Detection at Early Stage using Local Measures on MRI : A Comparative Study on Local Measures	224
<i>Sankara Nayaki K and Abraham Varghese</i>	
ECG Denoising using Adaptive Selection of IMFs through EMD and EEMD	228
<i>Gurpreet Singh, Gagandeep Kaur and Vineet Kumar</i>	

Classification based Land Use/Land Cover Change Detection through Landsat Images	232
--	-----

Rajeswari A V, Saritha S and G Santhosh Kumar

Author Index