DHEERAJ KUMAR

Assistant Professor,
Department of Electronics & Communication Engineering,
IIT Roorkee, Roorkee, Uttarakhand, India 247667

e-mail: dheeraj.kumar@ece.iitr.ac.in
genuine.dheeraj@gmail.com

Academic qualifications

Degree/Qualification	Year	Institute/ School
Ph.D. (Electrical & Electronic Engineering)	2012-2017	The University of Melbourne, Australia
B.Tech- M.Tech (Dual Degree) (Electrical Engineering)	2005-2010	Indian Institute of Technology (IIT) Kanpur, India

Scholastic achievements

- Selected for the prestigious **Ramanujan fellowship** offered by **Science and Engineering Research Board (SERB)**, Govt. of India
- One of 150 recipients of Melbourne International Fee Remission Scholarship (**MIFRS**) and Melbourne International Research Scholarship (**MIRS**) out of over 3000 applicants
- Secured All India Rank **156** in IIT JEE 2005 (Among the top **0.08%** of **200,000** students)
- Awarded **Merit Scholarship** by CBSE for outstanding performance (**All India Rank 91**) in AIEEE 2005
- Awarded the **Certificate of Merit** by CBSE for being among the top **0.1%** of all the qualified students in physics in class XII

Academic work experience

- **Organization:** Indian Institute of Tecnology (IIT) Roorkee **Place:** Roorkee, Uttarakhand, India

Designation: Assistant Professor **Period:** April 2019-Present

- Visual approaches for exploratory data analysis: A survey of VAT family of algorithms

- **Organization:** Purdue University (Supervisor: Prof. Satish Ukkusuri) **Place:** West Lafayette, Indiana, USA

Designation: Post-Doctoral Research Assistant **Period:** May 2017-April 2019

- Investigating the impacts of **mobility service providers** such as Uber on the **urban taxi market**
  - Collected and analyzed the trajectory data for Uber drivers by crawling the Uber web URL
  - Validated the **phantom driver** theory for the first time using **data driven analysis**
  - Studied the impact of **surge pricing** on customers' and drivers' behavior
  - A paper describing this work is currently under review at **IEEE transactions on Intelligent Transportation Systems (ITS)**
- Leveraging social media for better modelling of evacuation decisions for emergency events
  - Experimented on hurricane Sandy and Matthew Twitter data for **analyzing evacuation time and location of residents**
  - Analyzed the tweets to explore the **causation of their evacuation related decisions**

- **Organization:** RMIT University (Supervisor: Prof. Xiuzhen Zhang) **Place:** Melbourne, Australia

Designation: Research Officer **Period:** Oct 2016-March 2017

- Working on the problem of **Opinion spam detection** for online review sites
  - Proposed the use of an **inductive matrix completion** scheme for detecting **spammer groups of singleton reviewers**
  - Singleton reviewers are difficult to detect due to non-availability of spam indicator signals
  - Experimental results show the presence of several **singleton spammer groups** targeting various products on **iTunes and Amazon data**.

- **Organization:** The LNM Institute of Information Technology **Place:** Jaipur, India

Designation: Lecturer **Period:** Nov 2011-July 2012

- Taught **Microprocessor and Interface** using Intel's 8085 architecture and assembly programming as core of the course
- Set up **ATMEL MCU University center** at LNMIIT having facilities including AVR microcontrollers and necessary interfaces
- **Set up the lab** for Microprocessor, **designed the experiments** and carried them out successfully from scratch
- Member of equipment procuring committee and counseling cell

Teaching experience

- ECN 316: Digital Image Processing (IIT Roorkee)
- ECN 511: Linear Algebra and Random Processes (IIT Roorkee)
- CE 597: Data Science for Smart Cities (Purdue University)
- MICROI: Microprocessor and Interface (LNMIIT)

Publications

Book Chapter

● S. Mahallati, J.C. Bezdek, D. Kumar, M.R. Popovic, and T.A. Valiante, “Interpreting Cluster Structure in Waveform Data with Visual Assessment and Dunn's Index.” **Frontiers in Computational Intelligence - Springer**, pp. 73–101, 2017.

Journal publications

● P. Rathore, D. Kumar, S. Rajasegarar, M. S. Palaniswami, and J. C. Bezdek "Visual Structural Assessment and Anomaly Detection for High-Velocity Data Streams," in **IEEE Transactions on Cybernetics (T-CYB)**, accepted.

● X. Qian, D. Kumar, W. Zhang, and S. V. Ukkusuri, "Understanding the operational dynamics of Mobility Service Providers: A case of Uber," in **ACM Transactions on Spatial Algorithms and Systems (TSAS)**, vol. 6, no. 2, pp. 12:1-12:20, Feb. 2020.

● D. Kumar and J. C. Bezdek "Visual approaches for exploratory data analysis: A survey of the VAT family of algorithms," in **IEEE Systems, Man, and Cybernetics Magazine (SMC-MAG)**, accepted.

● M. Palaniswami, A. S. Rao, D. Kumar, P. Rathore, and S. Rajasegarar, "Role of Visual Assessment of Clusters for Big Data Analysis from Real-world Internet of Things," in **IEEE Systems, Man, and Cybernetics Magazine (SMC-MAG)**, accepted.

● P. Rathore, D. Kumar, S. Rajasegarar, M. S. Palaniswami, and J. C. Bezdek "A Scalable Framework for Trajectory Prediction," in **IEEE Transactions on Intelligent Transportation Systems (T-ITS)**, vol. 20, no. 10, pp. 3860-3874, Oct. 2019.

● P. Rathore, D. Kumar, J. C. Bezdek, S. Rajasegarar and M. S. Palaniswami, "A Rapid Hybrid Clustering Algorithm for Large Volumes of High Dimensional Data," in **IEEE Transactions on Knowledge & Data Engineering (TKDE)**, vol. 31, no. 4, pp. 641-654, Apr. 2019.

● D. Kumar, Z. Ghafoori, J. C. Bezdek, C. Leckie, K. Ramamohanarao, and M., Palaniswami, "Dealing with Inliers in Feature Vector Data," in **International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (IJUFKS)**, vol. 26, no. 2, pp. 25-45, 2018.

● D. Kumar, H. Wu, S. Rajasegarar, C. Leckie, S. Krishnaswamy and M. Palaniswami, "Fast and Scalable Big Data Trajectory Clustering for Understanding Urban Mobility," in **IEEE Transactions on Intelligent Transportation Systems (T-ITS)**, vol. 19, no. 11, pp. 3709-3722, Nov. 2018.

● P. Rathore, D. Kumar, S. Rajasegarar, and M. Palaniswami, “Maximum Entropy based Auto Drift Correction using High and Low Precision Sensors.” **ACM Transactions on Sensor Networks (TOSN)**, vol. 13, no. 3, pp. 24:1-24:41, Apr. 2017.

● D. Kumar, J. Bezdek, S. Rajasegarar, M. Palaniswami, C. Leckie, J. Chan, and J. Gubbi, “Adaptive Cluster Tendency Visualization and Anomaly Detection for Streaming Data.” **ACM Transactions on Knowledge Discovery from Data (TKDD)**, vol. 11, no. 2, pp. 24:1-24:40, Dec 2016.

● D. Kumar, J. Bezdek, S. Rajasegarar, C. Leckie, and M. Palaniswami, “A Visual-Numeric Approach to Clustering and Anomaly Detection for Trajectory Data.” **The Visual Computer - Springer**, vol. 33, no. 3, pp. 265-281, 2017.

● D. Kumar, J. Bezdek, M. Palaniswami, S. Rajasegarar, C. Leckie, and T. Havens, “A Hybrid Approach to Clustering in Big Data.” **IEEE Transactions on Cybernetics (T-CYB)**, vol. 46, no. 10, pp. 2372-2385, Oct. 2016.

● D. Kumar, S. Rajasegarar, and M. Palaniswami, “Geospatial estimation based auto drift correction in wireless sensor networks,” **ACM Transactions on Sensor Networks (TOSN)**, vol. 11, no. 3, pp. 50:1–50:39, Apr. 2015.

● D. Kumar, P. Vimal and Rajesh M. Hegde, “On the Soft Fusion of Probability Mass Functions for Multimodal Speech Processing,” **EURASIP Journal on Advances in Signal Processing**, vol. 2011, Article ID 294010

Conference Publications

● D. Kumar, T. Yabe, and S. Ukkusuri, “Social-Media aided Hyperlocal Help-Network Matching & Routing during Emergencies.” **IEEE International Conference on Big Data (BigData)**, pp. 1606-1611, 2018.

● P. Rathore, J. Bezdek, D. Kumar, S. Rajasegarar, and M. Palaniswami, “Approximate Cluster Heat Maps of Large High-Dimensional Data.” **International Conference on Pattern Recognition (ICPR)**, pp. 195-200, 2018.

● D. Kumar and S. Ukkusuri, “Utilizing Geo-tagged Tweets to understand Evacuation Dynamics during Emergencies: A case study of Hurricane Sandy.” **The Web Conference (WWW) Companion**, pp. 1613-1620, 2018.

● D. Kumar, Y. Shaalan, X. Zhang, and J. Chan, “Identifying Singleton Spammers via Spammer Group Detection.” **Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)**, pp. 656-667, 2018.

● P. Rathore, D. Kumar, S. Rajasegarar, and M. Palaniswami, “Bayesian Maximum Entropy and Interacting Multiple Model Based Automatic Sensor Drift Detection and Correction in an IoT Environment.” **IEEE World Forum on Internet of Things (WF-IoT)**, pp. 598-603, 2018.

● W. Zhang, D. Kumar, and S. V. Ukkusuri, “Exploring the Dynamics of Surge Pricing in Mobility-on-Demand Taxi Services,” **IEEE International Conference on Big Data (BigData)**, pp. 1375-1380, 2017.

● D. Kumar, H. Wu, Y. Lu, S. Krishnaswami, and M. Palaniswami, “Understanding Urban Mobility via Taxi Trip Clustering,” **IEEE International conference on Mobile Data Management (MDM)**, pp. 318-324, 2016.

● D. Kumar, S. Rajasegarar, M. Palaniswami, X. Wang, and C. Leckie, “A Scalable Framework for Clustering Vehicle Trajectories in a Dense Road Network,” **International Workshop on Urban Computing (UrbComp),** in conjunction with the **ACM SIGKDD** 2015.

● D. Kumar, J. Bezdek, S. Rajasegarar, M. Palaniswami, T. Havens, and C. Leckie, “clusiVAT: A mixed visual/numerical clustering algorithm for big data,” **IEEE International Conference on Big Data (BigData)**, pp. 112-117, 2013.

● D. Kumar, J. Gubbi, B. Yan, and M. Palaniswami, "Motor recovery monitoring in post-acute stroke patients using wireless accelerometer and cross correlation," **IEEE International Conference of the EMBS (EMBC)**, pp. 6703–6707, 2013.

● J. Gubbi, D. Kumar, A. Rao, B. Yan, and M. Palaniswami, "A pilot study on the use of accelerometer sensors for monitoring postacute stroke patients," **IEEE International Conference of the EMBS (EMBC)**, pp. 957–960, 2013.

● D. Kumar, S. Rajasegarar, and M. Palaniswami, "Automatic sensor drift detection and correction using spatial kriging and kalman filtering," **IEEE International Conference on Distributed Computing in Sensor Systems (DCoSS)**, pp. 183–190, 2013.

● D. Kumar, R. Malhotra, A. Singh and Rajesh M. Hegde, "Multimodal Speaker Diarization using a Soft Belief Function," **International Conference on Natural Language Processing (ICON)**, pp. 376-381, 2009.

Ph.D. Research (Sept 2012 – Oct 2016)

Supervisor: Prof. Marimuthu Palaniswami (Palani)

Adaptive cluster tendency visualization and anomaly detection for streaming data

- **Developed streaming data adaptation** of popular **cluster tendency assessment** algorithms for **static data**: VAT and iVAT
- The new algorithms **efficiently insert a new point** or **remove an existing point** from VAT **Minimum Spanning Tree (MST)**
- Demonstrated the applicability of new algorithms for **(visual) anomaly detection in evolving data streams** and for **sliding window based cluster assessment** for time series using **real life smart city IoT data**

Achievement: Faster algorithms (by several orders of magnitude) for cluster tendency assessment in streaming data

A mixed visual and numerical approach for big data clustering

- Designed a new single linkage based **clusiVAT** algorithm for **assessment of cluster tendency** in **Big Data**
- clusiVAT **samples the big data intelligently** and reorders the distance matrix to **estimate the clusters in the data visually**
- **Single linkage (SL)** is used to find clusters in the samples and the labels are extended to entire data using **nearest prototype rule**
- Experiments on real and synthetic datasets validated the **supremacy of clusiVAT over other Big Data clustering algorithms like k-means and CURE** in terms of partition accuracy and time taken

Achievement: Achieved **100% partition accuracy for 1 million 2D points in less than 2.5 sec**

A visual-numeric approach to clustering and anomaly detection for trajectory data

- Developed two **novel anomaly detection** approaches named **iVAT+** and **clusiVAT+** for small and big data case respectively
- Applied **iVAT+** to two publicly available **synthetic trajectory datasets** giving **better anomaly detection accuracy**
- Proposed a **novel two stage clusiVAT clustering on a real life MIT trajectories dataset** of vehicles and pedestrians from a parking lot scene

Achievement: More natural and informative trajectory clusters from MIT parking lot scenes

Immunizing single linkage clustering and iVAT visualization to inter cluster bridges

- **Proposed** two approaches **to make SL and VAT/iVAT robust to outliers** using the **Local Outlier Factor (LOF) score** of the points
- **Distance modification** approach **incorporated the LOF scores in distance calculation** to **avoid the chaining effect**
- **Data cleansing** approach **removes points that have an LOF score larger than some threshold** before applying VAT/iVAT
- Experiments on real life IoT data demonstrate that **while iVAT fails in the presence of anomalies**, the proposed new methods provides **correct assessment of cluster tendency**

Achievement: Proposed and validated a method to solve fundamental problem of **chaining in linkage based clustering**

Sensor drift detection for smart city applications

- Proposed a **framework to automatically detect and correct the drift** of a large number of inexpensive, error prone sensor nodes used for "**Smart City**" implementation
- Used **Kriging based interpolation** to predict actual value of the physical variable and **Kalman filter to get correct drift estimates**
- Demonstrate **practical usability** of this scheme in real time on a **Wireless Sensor Network (WSN)** using **Libelium Waspmote**
- Experimental demonstration on real sensor data obtained from **Intel Research Berkeley Laboratory deployment**
- Proposed and demonstrated two **power saving schemes** to prolong the usable lifecycle of the WSN

Achievement: Achieved **up to 8% reduction in sensor data error** as compared to **averaging based methods**

Motor recovery monitoring for post acute stroke patients

- Proposed an algorithm for **automatic stroke patient management** using a simple **hand wearable wireless system**
- Records activity of hands using **tri-axial accelerometer sensor**, and **predict NIHSS stroke index** based on activity comparison
- Captured hand's ability to perform different **rotational motion** using **cross correlation based features**

Achievement: Better accuracy for **prediction of NIHSS stroke index by up to 7%** from previously used **energy based methods**

Research Experience - External to the home Institution

- **Organization:** Institute for Infocomm Research (I2R) - A Star **Place:** Singapore
  Position: Foreign Student Attachment **Period:** 15th Nov, 2015 – 15th Feb, 2016

- Developed an algorithm to **cluster passenger taxi trip trajectories** to understand **urban traffic flow patterns**
- Proposed a **novel Dijkstra based Dynamic time warping distance measure, trajDTW** between two trajectories, which is suitable for **large numbers of overlapping trajectories in a dense road network**
- Performed numerical evaluation on a **large scale taxi trajectory dataset** consisting of **3.28 million passenger trips** from **15,061 taxis** during the period of **one month** within **Singapore**
- Analyzed the **distribution of taxi trips with time** to gain insights about how **traffic flow changes with time**, thus suggesting how the **frequency of public transport** should vary with time of day on different routes

- **Organization:** SoCioTal (European Union - EU FP7 Project), IotLab and IERC **Place:** Palic, Republic of Serbia
  Course: SenZations summer school on applications of IoT and WSN **Period:** 2nd – 6th Sept, 2013
  - **Worked in a group** of 6 for **designing a "Mood and Weather adjusted personalized music player"** using IoT infrastructure at **SmartSantandar** facility
  - Attended talks by eminent speakers in the field of IoT and Smart City implementation and applications
  - Learned about **leading European research projects and open challenges in the attractive area of M2M**
- **Organization:** Microsoft Research India **Place:** Indian Institute of Science, Bangalore, India
  Course: Summer School on "Computing for Socio-Economic development" **Period:** 13th – 17th June, 2010
  - Surveyed a few poor families and made a poster on "**How they manage their economics**"
  - Attended talks by eminent speakers in the field of **Information and Communication Technologies for Development**
  - Prepared a research proposal on "**Environmental impacts of ICT through e-governance**" with research focus on digitization of land records program in Karnataka, India, "BHOOMI"