Graph Mining and Analysis

……

Overview

Increasingly, today's massive data is in the form of complex graphs and networks. Examples include the world wide web, social networks, biological networks, semantic networks, and so on. The study of such complex and large scale networks (often referred to as network science) – understanding their intrinsic properties, changes to their structure over time or due to other external factors, and the behavior of entities and communities within them – can afford important insight to domain researchers and organizations alike. Given that networks are part and parcel of the complex and connected social, physical and biological world we live in, a coordinated and concerted approach combining these strands of research is essential to make progress on this grand challenge complex systems problem of our times.

In this course, we study the fundamental algorithms to model and mine graph data. We focus on graph and network modeling, graph pattern mining, as well as graph clustering and classification tasks. Course participants will learn these topics through lectures and hands-on tutorials. Also case studies and assignments will be shared to stimulate research motivation of participants.

The primary objectives of the course are as follows:

- Comprehensive understanding of the fundamental tasks and approaches in graph analytics and mining
- Hands-on tasks to solidify the graph mining concepts
- Appreciation of the cutting-edge research questions in graph analysis and mining

Lecture-wise course plan: (December 17-23, 2015, 24 Hours)

THURSDAY, DECEMBER 17, 2015:

- **Lecture 1:** 9:15 am – 10:30 am
  - Graph Analysis: Topological Properties and Centralities
- **Lecture 2:** 10:45 am – 12:00 noon
  - Graph Models: Heavy-tail degree distributions, Clustering effect, Small world, ER model, Strogatz-Watz Model, BA model, Kronecker Graphs, etc.
- **Hands-on Tutorial 1**: 12:30 pm – 1:30 pm

FRIDAY, DECEMBER 18, 2015:

- **Lecture 3:** 9:00am – 10:15am
  - Graph Pattern Mining: Frequent Subgraphs, Subgraph Isomorphism
- **Lecture 4:** 10:15am – 11:30am
  - Graph Mining Algorithms: Efficient Methods, Subgraph Sampling
- **Hands-on Tutorial 2:** 11:45am – 12:45pm

SATURDAY, DECEMBER 19, 2015:

- **Lecture 5**: 9:15 am – 10:30 am
  - Graph Clustering I: Spectral Clustering (Normalized Cut, Ratio Cut, Modularity, etc.), Girvan-Newman method, Clique percolation, etc.
- **Lecture 6**: 10:45 am – 12:00 noon
  - Graph Clustering II: Markov Chain Clustering, Hierarchical Clustering (Geodesic distance, average commute time), etc.
- **Hands-on Tutorial 3**: 12:30 pm – 1:30pm

SUNDAY, DECEMBER 20, 2015:

- **Lecture 7**: 9:15 am – 10:30 am
  - Graph Kernels: Kernel methods, Diffusion kernels (power kernel, exponential diffusion, von-neumann, etc.).
- **Lecture 8**: 10:45 am – 12:00 noon
  - Graph classification: SVMs, Pattern based kernels.
- **Hands-on Tutorial 4**: 12:30 pm – 1:30pm

MONDAY, DECEMBER 21, 2015:

- **Lecture 9**: 9:15 am – 10:30 am
  - Graph Classification II: Kernels between graphs (graphlets, random-walks, etc.)
- **Lecture 10**: 10:45 am – 12:00 noon
  - Graph Reachability: Fast reachability queries in directed graphs.
- **Hands-on Tutorial 5**: 12:30 pm – 1:30 pm

TUESDAY, DECEMBER 22, 2015:

- **Lecture 11**: 9:15 am – 10:30 am
  - Graph Matching Queries
- **Lecture 12**: 10:45 am – 12:00 noon
  - Graph Mining Summary: mining complex attributed graphs.
- **Hands-on Tutorial 6**: 12:30 pm – 1:30 pm

WEDNESDAY, DECEMBER 23, 2015:

- **Exam and Evaluation:** 10:00 am to 1:00 pm

Modules	**A: Duration:** **December 17 – 23, 2015 (24 Hours)** **B: Venue** **Department of Computer Science, Jamia Millia Islamia (A Central University), New Delhi – 25, India** ***Number of participants for the course will be limited to fifty.***
You Should Attend If…	▪ you are an executive, engineer and researcher from industry and government organizations, including R&D laboratories interested in graph analytics and mining. ▪ you are a student at all levels (BTech/MSc/MTech/PhD) or Faculty from reputed academic institutions interested in pursuing research career in data analytics and mining.
Fees	The participation fees for taking the course is as follows: **Participants from abroad : US $500** **Industry/ Research Organizations: INR 10000** **Academic Institutions:** • **Faculty members: Rs. 5000/-** • **Students: Rs. 2000/-** The above fee include all instructional materials, tutorials and assignments, laboratory equipment usage charges, 24 hour free internet facility. The participants will be provided accommodation on payment basis, subject to the availability.

The Faculty

Prof. Mohammed J. Zaki is a Professor of Computer Science at Rensselaer Polytechnic Institute (RPI) in Troy, NY, USA. He was also a Principal Scientist and Head of the Data Analytics group at the Qatar Computing Research Institute, Doha, Qatar from June 2013 to August 2015. His research interests focus on developing novel data mining techniques, especially for applications in bioinformatics and social networks. He has published over 225 papers and book-chapters on data mining and bioinformatics, including the recent textbook on Data Mining and Analysis: Fundamental Concepts and Algorithms (2014), published by Cambridge University Press.

He is the founding co-chair for the BIOKDD series of workshops. He is currently Area Editor for Statistical Analysis and Data Mining, and an Associate Editor for Data Mining and Knowledge Discovery, ACM Transactions on Knowledge Discovery from Data, and Social Networks and Mining. He received the National Science Foundation CAREER Award in 2001 and the Department of Energy Early Career Principal Investigator Award in 2002. He received an HP Innovation Research Award in 2010, 2011, and 2012, and a Google Faculty Research Award in 2011. He is a senior member of the IEEE, and an ACM Distinguished Scientist. His research is supported in part by NSF, NIH, DOE, Google, HP, and Nvidia. Further details about Prof. M. J. Zaki can be seen at his homepage: http://www.cs.rpi.edu/~zaki/

Dr. Muhammad Abulaish is an Associate Professor and Head of the Department of Computer Science, Jamia Millia Islamia, Delhi with over 17 years of experience in Academic and Research. He was also an Associate Professor and Head of the Internet Surveillance and Forensics research group at the Centre of Excellence in Information Assurance, King Saud University, Riyadh.

Abulaish's research interests are in the areas of Data Analytics and Mining, Social Computing, Cyber Forensics, Predictive Modeling, and Opinion Mining & Sentiment Analysis. He is keenly interested in developing analytical frameworks for integrated analysis of unstructured and structured data using data mining techniques for varied applications, including business intelligence, social network analysis, cyber forensics, open source intelligence, and web surveillance. Abulaish has over 75 research publications. He is a senior member of IEEE, ACM, and CSI. He is also a life member of ISTE, IETE, and ISCA. He is a program committee member of several international conferences, editorial board member/reviewer of various journals of repute in his area of research. Further details about Muhammad Abulaish can be seen at his homepage: http://jmi.ac.in/mabulaish

Course Co-ordinator

Dr. Muhammad Abulaish
Department of Computer Science
Jamia Millia Islamia (A Central University)
Jamia Nagar, New Delhi – 110025, India

Mobile: 07503607446
E-mail: mabulaish@jmi.ac.in

.....

Course Registration Link:

http://www.gian.iitkgp.ac.in/GREGN