

---

## Research Interests

Large-scale Data Integration, Fairness in Machine Learning and Software Testing and Social Network Analysis

---

## Education

- 2015-present **Master's and PhD in Computer Science**, *University of Massachusetts Amherst, USA*.  
Advisor: Prof. Barna Saha CGPA: 4.0/4.0
- 2010-2014 **BTech in Computer Science and Engineering**, *Indian Institute of Technology (IIT) Delhi, India*.  
Advisor: Prof. Amitabha Bagchi

---

## Selected Awards and Honors

- 2018 PhD Candidate with Distinction
- 2018 Finalist for Adobe Fellowship 2018
- 2017 Best paper award in SIGSOFT FSE 2017 and most reproducible award in SIGMOD 2017
- 2016 First recipient of Krithi Ramamritham Computer Science Scholarship (2016-17)
- 2016-18 Received SIGMOD travel award consecutively for three years
- 2014-18 Travel Grants: UAI 2014, AAAI 2018
- 2010 All India Rank 53 in IIT-JEE
- 2010 State Rank 1 and All India Rank 20 in AIEEE

---

## Selected Publications

- SIGMOD 2018 Robust Entity Resolution using Random Graphs.  
**Sainyam Galhotra**, Donatella Firmani, Barna Saha and Divesh Srivastava
- AAAI 2018, The Geometric Block Model
- Allerton 2018 **Sainyam Galhotra**, Arya Mazumdar, Soumyabrata Pal and Barna Saha
- ESEC/SIGSOFT Fairness Testing: Testing Software for Discrimination  
FSE 2017 **Sainyam Galhotra**, Yuriy Brun and Alexandra Meliou  
**Winner of Best paper award**
- SIGMOD 2017 Debunking the Myths of Influence Maximization  
**Sainyam Galhotra\***, Akhil Arora\* and Sayan Ranu  
**Winner of Most reproducible paper award**
- SIGMOD 2016 Holistic IM: Combining Scalability and Efficiency with Opinion-Aware Models  
**Sainyam Galhotra\***, Akhil Arora\* and Shourya Roy
- VLDB 2016 Tracking the Conductance of Rapidly Evolving Topic-Subgraphs  
**Sainyam Galhotra**, Amitabha Bagchi, Srikanta Bedathur, Maya Ramanath and Vidit Jain
- WWW 2015 ASIM: A Scalable Algorithm for Influence Maximization under the Independent Cascade Model  
**Sainyam Galhotra**, Akhil Arora, Srinivas Virinchi and Shourya Roy
- UAI 2014 Min-d-Occur: Ensuring Future Occurrences in Streaming Sets  
Vidit Jain and **Sainyam Galhotra**

---

## Work Experience

- May-Aug 2018 **Research Intern**, Megagon Labs, Mountain View, USA  
Mentor: Dr. Behzad Golshan, Dr. Wang-Chiew Tan  
Devised a holistic system to generate labelled training data with minimum human effort.
- May-Aug 2016 **Research Intern**, Google Research, Mountain View, USA  
Mentor: Dr. Steven Whang  
Devised efficient and scalable algorithms for recommending services to fasten the feature engineering pipeline on billions of internal datasets.
- 2014-2015 **Budding Scientist**, Xerox Research Center India (XRCI)  
Member of Text and Graph Analytics team  
Worked on devising scalable algorithms to solve complex problems in the area of data mining and graph mining.
- May-Aug 2013 **Research Intern**, Yahoo Labs, Bangalore, India  
Mentor: Dr. Vidit Jain  
Designed an algorithm to predict trending topics and tested it on real time twitter data. This work has been published in UAI - 2014.
- May-Aug 2012 **Research Intern**, École Nationale Supérieure des Télécommunications  
Mentor: Dr. Fabien Dagnat  
Corrected and improved the Continuations extension of Java Virtual Machine (JVM). Extended continuations to implement a prototype of Dynamic Software Updating.

---

## Professional Service

- 2015-2018 **Reviewer**: NIPS (2018, 2016), CIKM (2018, 2017), KDD (2018, 2017), SIGMOD 2018, TKDE, SEA 2017, WWW 2017, CoDS 2015
- 2018-2019 **PC Member**: NDA Workshop SIGMOD 2018 and 2019, SIGMOD Reproducibility 2018
- 2016-2017 **Teaching Assistant**: Algorithms for data science (CS590D) at UMass Amherst
- 2018-19 **Graduate Student Representative**
- 2015-18 Mentored 4 BTech students for their undergraduate thesis at IIT Delhi
- 2017 **Student Mentor** for 2 incoming graduate students at UMass Amherst
- 2015 Grader for a course namely, (Inside the Box) CS335 at UMass Amherst
- 2015 Intern Mentor at Xerox Research Centre India (XRCI) for Srinivas Virinchi, PhD student (UMD, College Park)
- 2015 Organizing Committee: XRCI Open

---

## Preprints

- P1 Afro: Progressive End-to-End Entity Resolution  
**Sainyam Galhotra**, Donatella Firmani, Barna Saha and Divesh Srivastava
- P2 Connectivity in Random Annulus Graphs and the Geometric Block Model  
**Sainyam Galhotra**, Soumyabrata Pal, Arya Mazumdar and Barna Saha
- P3 Explainable Labelled Data Generation  
**Sainyam Galhotra**, Behzad Golshan and Wang-Chiew Tan

---

## Patents

- 2015 System to predict the channel of customer contact  
**Sainyam Galhotra** and Narayanan Unny
- 2015 Methods and Systems for Identifying Target Users of Content  
Akhil Arora, **Sainyam Galhotra**, Srinivas Virinchi and Shourya Roy

- 2015 System for Identifying Root Causes of Churn for Churn Prediction Refinement  
Akhil Arora, Manoj Gupta, Neeta Pande, **Sainyam Galhotra** and Shourya Roy

---

## Journal Publications

- IEEE Data Eng. Bulletin Robust Entity Resolution Using a CrowdOracle  
Donatella Firmani, **Sainyam Galhotra**, Barna Saha, Divesh Srivastava
- TOSN Optimal Radius for Connectivity in Duty-Cycled Wireless Sensor Networks  
Amitabha Bagchi, Cristina M. Pinotti, **Sainyam Galhotra** and Tarun Mangla
- Pure and Applied Mathematics Journal An adaptation method for removing arsenate species from water solution  
**Sainyam Galhotra**, Shigeru Kanemitsu, Hiroyuki Kondo
- JCP Turing - Hopf instabilities through a combination of diffusion, advection and finite size effects  
**Sainyam Galhotra**, J. K. Bhattacharjee and Bijay K. Agarwalla
- JOMC Diffusion driven instability to a drift driven one: Turing patterns in the presence of an electric field  
Bijay K. Agarwalla, **Sainyam Galhotra** and J. K. Bhattacharjee

---

## Other Publications

- CoDS/COMAD 2018 Influence Maximization Revisited: The State of the Art and the Gaps that Remain  
Tutorial **Sainyam Galhotra\***, Akhil Arora\* and Sayan Ranu
- NEDB 2016 Debunking the Myths of Influence Maximization  
(Oral) **Sainyam Galhotra\***, Akhil Arora\* and Sayan Ranu
- AAAI 2016 QA<sup>RT</sup> : A System for Real-time Holistic Quality Assurance for Contact Center Dialogues  
Shourya Roy, Sandipan Dandapat, R. Mariappan, S Srivastava, **Sainyam Galhotra** and B. Peddamuthu
- NEDB 2016 Holistic IM: Combining Scalability and Efficiency with Opinion-Aware Models  
(Poster) **Sainyam Galhotra\***, Akhil Arora\* and Shourya Roy
- CODS/COMAD 2015 STAR: Real-time Spatio-Temporal Analysis and Prediction of Traffic Insights using Social Media  
Deepali Semwal, Sonal Patil, **Sainyam Galhotra**, Akhil Arora and Narayanan Unny
- MSWIM 2013 Optimal Radius for Connectivity in Duty-Cycled Wireless Sensor Networks  
Amitabha Bagchi, Cristina M. Pinotti, **Sainyam Galhotra** and Tarun Mangla
- SUDA 2013 An adaptation method for removing arsenate species from water solution  
**Sainyam Galhotra** and Shigeru Kanemitsu

---

## Talks

- Sept 2018 *Local Correlation Clustering*, UMass Theory seminar
- June 2018 *Repairing Noisy Graphs*, Megagon Labs
- Jan 2018 *Influence Maximization Revisited: The Gaps that Remain*, CODS/COMAD 2018
- Jan 2017 *Debunking the Myths of Influence Maximization*, NEDB
- Sept 2016 *Tracking the Conductance of Rapidly Evolving Topic-Subgraphs*, IIT Delhi
- July 2016 *Tracking the Conductance of Rapidly Evolving Topic-Subgraphs*, NDA Workshop
- June 2016 *Holistic IM: Combining Scalability and Efficiency with Opinion-Aware Models*, UCSB
- June 2016 *Holistic IM: Combining Scalability and Efficiency with Opinion-Aware Models*, Facebook Inc
- June 2016 *Holistic IM: Combining Scalability and Efficiency with Opinion-Aware Models*, PARC
- May 2015 *ASIM: A Scalable Algorithm for Influence Maximization under the Independent Cascade Model*, Univ. of Perugia Italy

---

## Technical Skills

**Programming languages:** C, C++, Python, Java, R, Matlab, SQL

**Packages used:** Numpy, Scipy, Boost, OpenGL

**Applications and Tools:** L<sup>A</sup>T<sub>E</sub>X, Matlab, GNU Plot, Shell Scripting, OpenOffice, MS Office

**Worked with large graph datasets:** NetHEPT (15K nodes and 62K edges), DBLP (317K nodes and 1M edges), YouTube (1M nodes and 3M edges), Twitter (7M nodes and 3B edges) and Friendster (65M nodes and 1.8B edges)

---

## Positions of Responsibility

2017-18 **Vice President, Indian Student Association, UMass Amherst**

- Helped incoming Indian students by providing airport pickup and temporary accommodation.
- Organized cultural programs to celebrate Indian festivals like Diwali and Holi.

2014-15 **Overall Coordinator, XRCI-CSR (Corporate Social Responsibility)**

- Initiated the CSR group at Xerox Research Centre India and led the team from the front for the social activities like cleanliness campaign, distributing food to the needy and so on.

2014-15 **Committee Member, XRCI Open**

- Organised technical events at the conference at Xerox Research Centre India, Bangalore.