

Charu C. Aggarwal

Work Address

Charu C. Aggarwal
1101 Kitchawan Road,
Yorktown, NY 10598
Phone: (914) 602 8152 (Mobile)
Email: CharuCAggarwal@gmail.com

Personal Address

Charu C. Aggarwal
182 Scenic Drive,
Mohegan Lake, NY 10547
Phone: (914) 743 1423 (Residence)
Email: CharuCAggarwal@gmail.com

Professional Interest

- Specializations: Data mining, data streams, scalable data analytics, outlier analysis, social network analysis, recommender systems

Biography

Charu Aggarwal is a Distinguished Research Staff Member (DRSM) at the IBM T. J. Watson Research Center in Yorktown Heights, New York. He completed his B.Tech. from IIT Kanpur in 1993 and his Ph.D. from Massachusetts Institute of Technology in 1996. His research interests include data mining, with an emphasis in data streams and scalable data analytics. He has published over 300 papers, authored 4 books, edited 11 books, and has applied for or been granted over 80 patents. According to *Google Scholar*, his h-index is 73. Because of the commercial value of the aforementioned patents, he has received several invention achievement awards and has thrice been designated a Master Inventor by IBM. He is a recipient of an *IBM Corporate Award* (2003) for his work on bio-terrorist threat detection in data streams, a recipient of the *IBM Outstanding Innovation Award* (2008) for his scientific contributions to privacy technology, and a recipient of two *IBM Outstanding Technical achievement Awards* (2009, 2016) for his scientific contributions to data streams and high-dimensional data. He also received the *EDBT 2014 test-of-time award* for the long-term impact of one of his results on privacy. He served as an associate editor of the *IEEE Transactions on Knowledge and Data Engineering* from 2004 to 2008. He is an associate editor of the *ACM Transactions on Knowledge Discovery and Data Mining*, an action editor of the *Data Mining and Knowledge Discovery Journal*, and an associate editor of the *Knowledge and Information Systems Journal*. He served as program co-chair/general co-chair of IEEE Big Data Conference (2014), IEEE ICDM Conference (2015), CIKM Conference (2015), ACM KDD Conference (2016), and is editor-in-chief of the *ACM SIGKDD Explorations*. He is a fellow of the ACM, SIAM, and the IEEE for contributions to knowledge discovery and data mining algorithms. He is a recipient of the IEEE ICDM Research Contributions Award (2015), which is one of two highest awards for research contributions in the field of data mining.

Education

Massachusetts Institute of Technology

Cambridge, MA

- Ph.D. in Operations Research, April 1996.
- **Thesis Title:** Faster Algorithms for Some Network Flow Problems.
- **Thesis Advisor:** Professor James B. Orlin

Indian Institute of Technology, Kanpur

Kanpur, India

- Bachelor of Technology in Computer Science and Engineering, May 1993.
- Graduated with grade point average of 10.0 out of 10.0
- Received award for best finishing undergraduate in the department of Computer Science at IIT Kanpur.

Experience

IBM T. J. Watson Research Center

Yorktown Heights, NY (1996-present)

- Research staff member at IBM Research (1996-2014).
- Distinguished Research Staff Member at IBM (2014-present).
- Primary Interests in streaming data, large scale graphs and social networks, and privacy/uncertainty issues with scale.
- Designed algorithms for large scale and streaming data analytics, streaming uncertain data, privacy, high dimensional data analytics and text data.
- Broader Interests in data mining, business intelligence, social networks, privacy, and data streams.

Massachusetts Institute of Technology

Cambridge, MA (1993- 1996)

- Research and teaching assistant
- Research in Network Flows, Dynamic Data Structures, and Genetic Algorithms for Combinatorial Optimization
- Twice served as teaching assistant for the course 15.082 (Network Flow Algorithms)

Honors

- **IEEE ICDM Research Contributions Award**, which is one of two highest awards for research achievements in the field of data mining.
- **SIAM Fellow** for “*contribution to knowledge discovery and data mining algorithms.*”, 2015.
- Invited Keynote Speaker, ECML Conference, 2014.
- Invited Keynote Speaker, ASONAM Conference, 2014.
- **ACM Fellow** for “*contribution to knowledge discovery and data mining algorithms.*”, 2013.
- **EDBT 2014 Test of Time Award**, EDBT Conference, 2014.
- **Best Paper Award**, ASONAM Conference, 2011.
- **IEEE Fellow** for “*contributions to knowledge discovery and data mining algorithms.*”.
- **IBM Outstanding Technical Achievement Award** for contributions to research in high-dimensional data, 2016.
- **IBM Outstanding Technical Achievement Award** for contributions to System S (IBM Streams Product), 2010.
- **IBM Research Division Award** for contributions to stream mining in System S, 2008.
- **IBM Outstanding Innovation Award** for scientific contributions to privacy technology, 2008.
- **IBM Corporate Award for Environmental Excellence**, for Bio-terrorist threat detection from high dimensional data streams of medical and pharmaceutical records, 2003.
- Designated **Master Inventor** at IBM Research for commercial value of patents (2000-2003, 2003-2006, 2006-2010).
- Invited Keynote Speaker, European Conference on Machine Learning, 2006.
- IBM Twentieth Plateau Invention Achievement Award, (2013)
- IBM Nineteenth Plateau Invention Achievement Award, (2011)
- IBM Eighteenth Plateau Invention Achievement Award (2009)
- IBM Seventeenth Plateau Invention Achievement Award (2008)
- IBM Sixteenth Plateau Invention Achievement Award (2007)

- IBM Fifteenth Plateau Invention Achievement Award, (2006)
- IBM Fourteenth Plateau Invention Achievement Award (2005)
- IBM Thirteenth Plateau Invention Achievement Award (2004)
- IBM Twelfth Plateau Invention Achievement Award (2004)
- IBM Eleventh Plateau Invention Achievement Award (2003)
- IBM Tenth Plateau Invention Achievement Award (2002)
- IBM Ninth Plateau Invention Achievement Award (2001)
- IBM Eighth Plateau Invention Achievement Award (2000)
- IBM Seventh Plateau Invention Achievement Award (2000)
- IBM Sixth Plateau Invention Achievement Award. (1999)
- IBM Fifth Plateau Invention Achievement Award. (1998)
- IBM Fourth Plateau Invention Achievement Award. (1998)
- IBM Third Plateau Invention Achievement Award. (1997)
- IBM Second Plateau Invention Achievement Award. (1997)
- IBM First Plateau Invention Achievement Award. (1996)
- Sloan School of Management Fellowship, Massachusetts Institute of Technology (1995).
- Indian Institute of Technology award for best finishing undergraduate in the Department of Computer Science (1993)
- 1991 first prize for academic excellence (IIT Kanpur).
- 1990-1991 academic proficiency award (IIT Kanpur).
- 1989-1990 academic proficiency award (IIT Kanpur).

Representative Publications

The following is a list of representative publications, which is selected both on a combination of recent research interest, and historical impact. For publications authored more than five years earlier, an approximate citation count according to *Google Scholar* is provided:

1. **(Textbook)** Charu C. Aggarwal. Recommender Systems: The Textbook, Springer, 2016.
2. **(Textbook)** Charu C. Aggarwal. Data Mining: The Textbook, Springer, 2015.
3. **(Monograph, Textbook)** Charu C. Aggarwal. Outlier Analysis, Springer, 2013 (first edition is a monograph, and second edition (2017) is a textbook).
4. **(Edited Book)** Charu C. Aggarwal. Data Streams: Models and Algorithms, Springer, 2007. (> 650 citations)
5. Charu C. Aggarwal, Jiawei Han, Jianyong Wang, Philip Yu. A Framework for Clustering Evolving Data Streams. *Very Large Databases (VLDB) Conference*, pp. 81–92, 2003. (> 1500 citations)
6. Charu C. Aggarwal, Cecilia Procopiuc, Joel L. Wolf, Philip S. Yu, Jong Soo Park. Fast Algorithms for Projected Clustering. *ACM SIGMOD Conference on Management of Data*, pp. 61–72, 1999. (> 1000 citations)
7. Charu C. Aggarwal, Jiawei Han, Jianyong Wang, Philip S. Yu. A Framework For Projected Clustering of High Dimensional Data Streams. *Very Large Databases (VLDB) Conference*, pp. 852–863, 2004. (> 500 citations)
8. Dakshi Agrawal, Charu C. Aggarwal. On the Design and Quantification of Privacy Preserving Data Mining Algorithms. *ACM Principles on Databases Systems (PODS Conference)*, pp. 247–255, 2001. (> 1000 citations)

9. Charu C. Aggarwal. On Unifying Privacy and Uncertain Data Models. *IEEE International Conference on Data Engineering (ICDE)*, 2008.
10. Charu C. Aggarwal. On k -anonymity and the curse of dimensionality. *Very Large Databases (VLDB) Conference*, pp. 901–909, 2005. (> 500 citations)
11. Charu C. Aggarwal, Alexander Hinneburg, Daniel A. Keim. On the surprising behavior of distance metrics in high dimensional space. *International Conference on Database Theory (ICDT Conference)*, pp. 420–434, 2001. (> 900 citations)
12. Alexander Hinneburg, Charu C. Aggarwal, Daniel A. Keim. What is the nearest neighbor in high dimensional space? *Very Large Databases (VLDB) Conference*, pp. 506–515, 2000. (> 500 citations)
13. Charu C. Aggarwal, Karthik Subbian. Event Detection in Social Streams, *SIAM Conference on Data Mining (SDM)*, pp. 624–635, 2012.
14. Charu C. Aggarwal, Philip S. Yu. Outlier Detection for High Dimensional Data. *ACM SIGMOD Conference on Management of Data*, pp. 37–46, 2001. (> 1000 citations)
15. Charu C. Aggarwal, Philip S. Yu. A Condensation Approach to Privacy Preserving Data Mining. *Conference on Extending Database Technology (EDBT)*, pp. 183–199, 2004. (> 350 citations) **EDBT 2014 Test-of-time award**
16. Ramesh C. Agarwal, Charu C. Aggarwal, V. V. V. Prasad. A Tree Projection Algorithm for Generation of Frequent Itemsets. *Journal of Parallel and Distributed computing*. 61(3), pp. 360–371, 2001. (> 750 citations)
17. Charu C. Aggarwal, Philip S. Yu. Finding Generalized Projected Clusters in High dimensional spaces. *ACM SIGMOD Conference on Management of Data*, pp. 70–81, 2000. (> 600 citations)

Citations

Google Scholar citations to my articles (PDF Document Clickable Link):
http://scholar.google.com/citations?user=x_wsduUAAAAJ&hl=en

Authored Textbooks (for Classroom Teaching)

1. Charu C. Aggarwal. Outlier Analysis, *Second Edition*, Springer, to appear, 2017 [The first edition (see below) from 2013 was officially classified as a monograph and was converted to a textbook at the suggestion of the publisher]
2. Charu C. Aggarwal. Recommender Systems: The Textbook, Springer, 2016.
3. Charu C. Aggarwal. Data Mining: The Textbook, Springer, 2016. (Features in Springer’s list of their top-20 best sellers in Computer Science)

Authored Books (for Researchers)

1. Charu C. Aggarwal and Saket Sathe. Outlier Ensembles: An Introduction, Springer, 2017 (to appear).
2. Manish Gupta, Jing Gao, Charu C. Aggarwal and Jiawei Han. Outlier Detection for Temporal Data, *Morgan and Claypool*, 2014.
3. Charu C. Aggarwal. Outlier Analysis, *First Edition*, Springer, 2013. (This book received a designation for being among the “Best publications of 2013” by ACM Computing Reviews. The book features consistently in Springer’s list of their top-100 best sellers in Computer Science. The second edition (see above) is being published as a textbook at the specific request of the publisher.)

Edited Books

1. Charu C. Aggarwal (editor). Data Classification: Algorithms and Applications, *CRC Press*, 2014.
2. Charu C. Aggarwal, Jiawei Han (editors) Frequent Pattern Mining, *Springer*, 2014.
3. Chandan K. Reddy, Charu C. Aggarwal (editors) Healthcare Data Analytics, *CRC Press*, 2015.
4. Charu C. Aggarwal, Chandan Reddy (Editors). Data Clustering: Algorithms and Applications, *CRC Press*, August 2013.
5. Charu C. Aggarwal (Editor). Managing and Mining Sensor Data, *Springer*, January 2013.
6. Charu C. Aggarwal, ChengXiang Zhai (Editors). Mining Text Data, *Springer*, 2012.
7. Charu C. Aggarwal (Editor). Social Network Data Analytics, *Springer*, 2011.
8. Charu C. Aggarwal, Haixun Wang (Editors). Managing and Mining Graph Data, *Springer*, 2010.
9. Charu C. Aggarwal (Editor). Managing and Mining Uncertain Data, *Springer*, 2009.
10. Charu C. Aggarwal, Philip S. Yu (Editors). Privacy-Preserving Data Mining: Models and Algorithms, Springer, 2008.
11. Charu C. Aggarwal (Editor). Data Streams: Models and Algorithms, Springer, 2007.

Journal Publications

1. Jiliang Tang, Yi Chang, Charu C. Aggarwal, and Huan Liu. A Survey of Signed Network Mining in Social Media. *ACM Computing Surveys*, accepted, 2016.
2. Guojun Qi, Wei Liu, Charu Aggarwal, and Thomas Huang. Joint Intermodal and Intramodel Transfers for Extremely Rare or Unseen Classes. *IEEE Transactions on Pattern Mining and Machine Intelligence*, accepted, 2016.
3. Karthik Subbian, Charu C. Aggarwal, and Jaideep Srivastava. Mining Influencers Using Information Flows in Social Streams. *ACM Transactions on Knowledge Discovery from Data*, 10(3), 26, 2016.
4. Hessam Zakerzadeh, Charu C. Aggarwal and Ken Barker. Managing dimensionality in data privacy anonymization. *Knowledge and Information Systems Journal*, accepted to appear, 2015.
5. Charu C. Aggarwal, Yao Li, and Philip S. Yu. On the anonymizability of graphs. *Knowledge and Information Systems*, 45(3), pp. 571–588, 2015.
6. Charu C. Aggarwal and Saket Sathe. Theoretical Foundations and Algorithms for Outlier Ensembles. *ACM SIGKDD Explorations*, 17(1), pp. 24–47, 2015.
7. Guo-Jun Qi, Charu C. Aggarwal, and Thomas S. Huang. Breaking the Barrier to Transferring Link Information across Networks. *IEEE Transactions on Knowledge and Data Engineering*, 27(7), pp. 1741–1753, 2015.
8. Charu C. Aggarwal and Karthik Subbian. Evolutionary Network Analysis: A Survey. *ACM Computing Surveys*, 47(1): 10:1-10:36, 2014.
9. Manish Gupta, Jing Gao, Charu C. Aggarwal, and Jiawei Han. Outlier Detection for Temporal Data: A Survey. *IEEE Transactions on Knowledge and Data Engineering*, 26(9), pp. 2250–2267, 2014.
10. Charu C. Aggarwal, Yuchen Zhao, and Philip S. Yu. On the use of side-information for mining text data, *IEEE Transactions on Knowledge and Data Engineering*, 26(6), pp. 1415–1429, 2014. (**Special Issue on “Best of IEEE ICDE, 2012”**)
11. Charu C. Aggarwal, Yan Xie, and Philip S. Yu. A Framework for Dynamic Link Prediction in Heterogeneous Networks, *Statistical Analysis and Data Mining*, 7(1): pp. 14–33, 2014. (**Special Issue on “Best of SIAM Conference On Data Mining, 2012”**)

12. Dong Wang, Lance M. Kaplan, Tarek F. Abdelzaher, Charu C. Aggarwal. On Credibility Estimation Tradeoffs in Assured Social Sensing. *IEEE Journal on Selected Areas in Communications*, 31(6), pp. 1026–1037, 2013.
13. Charu C. Aggarwal. On the Analytical Properties of High Dimensional Randomization, *IEEE Transactions on Knowledge and Data Engineering*, 25(7), pp. 1628–1642, 2013.
14. Charu C. Aggarwal, Peixiang Zhao. Towards graphical models for text processing. *Knowledge and Information Systems*, 36(1), pp. 1–21, 2013.
15. Mohammad M. Masud, Qing Chen, Latifur Khan, Charu C. Aggarwal, Jing Gao, Jiawei Han, Ashok N. Srivastava, Nikunj C. Oza. Classification and Adaptive Novel Class Detection of Feature-Evolving Data Streams. *IEEE Transactions on Knowledge and Data Engineering*, 25(7), pp. 1484–1497, 2013.
16. Charu C. Aggarwal. Mining Text and Social Streams: A Review. *ACM SIGKDD Explorations*, 15(2), pp. 9–19, 2013.
17. Charu C. Aggarwal, Nan Li. On Supervised Mining of Dynamic Content-based Networks, *Journal of Statistical Analysis and Data Mining*, 5(1), pp. 16–34, 2012. (**Special Issue on “Best of SIAM Conference on Data Mining, 2011”.**)
18. Guojun Qi, Charu C. Aggarwal, Qi Tian, Heng Ji, Thomas S. Huang. Exploring Content and Context Links in Social Media: A Latent Space Method. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 34(5), pp. 850–862, 2012.
19. Charu C. Aggarwal. A segment-based framework for modeling and mining data streams. *Knowledge and Information Systems (KAIS) Journal*, 30(1), pp. 1–29, 2012.
20. Yizhou Sun, Charu C. Aggarwal, Jiawei Han. Relation Strength-Aware Clustering of Heterogeneous Information Networks with Incomplete Attributes. *PVLDB Journal*, 5(5), pp. 394–405, 2012.
21. Charu C. Aggarwal. Outlier Ensembles: Position Paper, *ACM SIGKDD Explorations*, 14(2), pp. 49–58, 2012.
22. Charu C. Aggarwal: On the equivalence of PLSI and projected clustering. *SIGMOD Record*, 41(4), pp. 45–50, 2012.
23. Charu C. Aggarwal, Philip S. Yu. On the Network Effect in Web 2.0 Applications, *Electronic Commerce and Applications*, 11(2), pp. 142–152, 2012.
24. Fei Wang, Hanghang Tong, Philip S. Yu, and Charu C. Aggarwal. Guest editorial: special issue on data mining technologies for computational social science. *Data Mining and Knowledge Discovery*, 25(3), pp. 415–419, 2012.
25. Peixiang Zhao, Charu C. Aggarwal, Min Wang. gSketch: On Query Estimation in Graph Streams. *PVLDB Journal*, 5(3), pp. 193–204, 2011.
26. Charu C. Aggarwal, Chen Chen, Jiawei Han. The Inverse Classification Problem. *Journal of Computer Science and Technology*, 25(3), pp. 458–468, 2010.
27. Charu C. Aggarwal, Philip S. Yu. On clustering massive text and categorical data streams. *Knowledge and Information Systems (KAIS) Journal*, 24(2), pp. 171–196, 2010.
28. Charu C. Aggarwal, Yao Li, Philip S. Yu, Ruoming Jin. On Dense Pattern Mining in Graph Streams. *PVLDB Journal*, 3(1): pp. 975–984, 2010.
29. Charu C. Aggarwal, Yuchen Zhao, Philip S. Yu. A framework for clustering massive graph streams. *Statistical Analysis and Data Mining Journal*, 3(6), pp. 399–416, 2010. **Special Issue for “Best of SIAM Conference on Data Mining, 2010.”**
30. Charu C. Aggarwal, Yan Xie, Philip S. Yu. GConnect: A Connectivity Index for Massive Disk-resident Graphs. *PVLDB Journal*, 2(1), pp. 862–873, 2009.

31. Charu C. Aggarwal, On Classification and Segmentation of Massive Audio Data Streams, *Knowledge and Information Systems*, 20(2), pp. 137–156, 2009.
32. Charu C. Aggarwal, Philip S. Yu. A Survey of Uncertain Data Algorithms and Applications, *IEEE Transactions on Knowledge and Data Engineering*, 21(5), pp. 609–623, 2009.
33. Jianyong Wang, Youzhu Zhang, Lizhu Zhou, George Karypis, Charu C. Aggarwal. CONTOUR: An Efficient Algorithm for Discovering Discriminating Subsequences, *Data Mining and Knowledge Discovery Journal*, 18(1), pp. 1–29, Feb 2009.
34. Charu C. Aggarwal, Philip S. Yu. A Framework for Condensation-based Anonymization of String Data, *Data Mining and Knowledge Discovery Journal*, 16(3), pp. 251–275, July 2008.
35. Charu C. Aggarwal, Philip S. Yu. On Static and Dynamic Methods for Condensation-based Privacy-Preserving Data Mining. *ACM Transactions on Database Systems (ACM TODS Journal)*, 33(1), March 2008.
36. Charu C. Aggarwal. Towards Exploratory Test Instance Centered Diagnosis in High Dimensional Classification. *IEEE Transactions on Knowledge and Data Engineering*, 19(8), pp. 1001–1015, August, 2007.
37. Mohammed J. Zaki, Charu C. Aggarwal: XRules: An effective algorithm for structural classification of XML data. *Machine Learning Journal* 62(1-2): 137–170, 2006.
38. Charu C. Aggarwal, Jiawei Han, Jianyong Wang, Philip S. Yu. A Framework for On Demand Classification of Evolving Data Streams. *IEEE Transaction on Knowledge and Data Engineering*, 18(5), pp. 577–589, 2006.
39. Charu C. Aggarwal. On the use of human-computer interaction for meaningful high-dimensional nearest neighbor search. *Data Mining and Knowledge Discovery Journal*, 13(1), pp. 89–117, 2006.
40. Charu C. Aggarwal, Jiawei Han, Jianyong Wang, Philip S. Yu. On High Dimensional Projected Clustering of Data Streams. *Data Mining and Knowledge Discovery (DMKD) Journal*, 10(3), pp. 251–273, 2005.
41. Charu C. Aggarwal, Philip S. Yu. An efficient and effective algorithm for high dimensional outlier detection. *VLDB Journal*, 14(2), pp. 211–221, 2005.
42. Charu C. Aggarwal. On the use of Wavelet Decomposition for String Classification, *Data Mining and Knowledge Discovery (DMKD) Journal*, 10(2), pp. 117–139, 2005.
43. Charu C. Aggarwal. On Change Diagnosis in Evolving Data Streams. *IEEE Transactions on Knowledge and Data Engineering*, 17(5), pp. 587–600, 2005.
44. Charu C. Aggarwal. On Leveraging User Access Patterns for Topical Specific Crawling. *Data Mining and Knowledge Discovery (DMKD) Journal*, 9(2), pp. 123–146, 2004.
45. Charu C. Aggarwal. An Efficient Subspace Sampling Framework for High Dimensional Data Reduction, Selectivity Estimation, and Nearest Neighbor Search. *IEEE Transactions on Knowledge and Data Engineering*, 16(10), pp. 1247–1262, 2004.
46. Charu C. Aggarwal. A Human Computer Interactive Method for Projected Clustering. *IEEE Transactions on Knowledge and Data Engineering*, 16(4), pp. 448–460, 2004.
47. Charu C. Aggarwal, Stephen C. Gates, Philip S. Yu. On Using Partial Supervision for Text Categorization. *IEEE Transactions on Knowledge and Data Engineering* 16(2) pp. 245–255, 2004.
48. Srinivasan Parthasarathy, Charu C. Aggarwal. On the Use of Conceptual Reconstruction for Mining Massively Incomplete Data Sets. *IEEE Transactions on Knowledge and Data Engineering*, 15(6), pp. 1512–1521, 2003.
49. Charu C. Aggarwal. Towards Effective and Interpretable Data Mining by Visual Interaction. *ACM SIGKDD Explorations*, 3(2), pp. 11–22, 2002.

50. Charu C. Aggarwal, James B. Orlin. On Multiroute Maximum Flows in Networks. *Networks*, 39(1), pp. 43–52, 2002.
51. Charu C. Aggarwal, Zheng Sun, Philip S. Yu. Fast Algorithms for Online Generation of Profile Association Rules. *IEEE Transactions on Knowledge and Data Engineering*, 14(5), pp. 1017–1028, 2002.
52. Charu C. Aggarwal, Philip S. Yu. Redefining clustering for high dimensional applications. *IEEE Transactions on Knowledge and Data Engineering*, 14(2), pp. 210–225, 2002.
53. Charu C. Aggarwal, Cecilia Procopiuc, Philip S. Yu. Finding Localized Associations in Market Basket Data. *IEEE Transactions on Knowledge and Data Engineering*, 14(1), pp. 51–62, 2002.
54. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. Adaptive Piggybacking Schemes for Video-on-demand Systems. *Multimedia Tools and Applications*, 16(3), pp. 231–250, 2002.
55. Charu C. Aggarwal. Re-designing distance functions and distance based applications for high dimensional data. *ACM SIGMOD Record*, 30(1), pp. 13–18, March 2001.
56. Charu C. Aggarwal, Fatima Al-Garawi, Philip Yu. On the Design of a Learning Crawler for Topical Resource Discovery. *ACM Transactions on Information Systems*, 19(3), pp. 286–309, 2001.
57. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. The Maximum Factor Queue Length Batching Scheme for Video On Demand Systems. *IEEE Transactions on Computers*, 50(2), pp. 97–110, 2001.
58. Charu C. Aggarwal, Philip S. Yu. A New Approach to Online algorithms for generation of association rules. *IEEE Transactions on Knowledge and Data Engineering*, 13(4), pp. 527–540, 2001.
59. Charu C. Aggarwal, Philip S. Yu. Mining Associations with the Collective Strength Approach. *IEEE Transactions on Knowledge and Data Engineering*, 13(6), pp. 863–873, 2001.
60. Ramesh C. Agarwal, Charu C. Aggarwal, V. V. V. Prasad. A Tree Projection Algorithm for Generation of Frequent Itemsets. *Journal of Parallel and Distributed computing*. 61(3), pp. 360–371, 2001.
61. Charu C. Aggarwal, Philip S. Yu. Data Mining Techniques for Personalization. *IEEE Data Engineering Bulletin*, 23(1), pp. 4–9, 2000.
62. Charu C. Aggarwal, Joel L. Wolf, and Philip S. Yu. Design and Analysis of Permutation-Based Pyramid Broadcasting. *Multimedia Systems Journal*, 7(6), pp. 439–448, 1999.
63. Charu C. Aggarwal, Joel L. Wolf, and Philip S. Yu. Caching on the world wide web. *IEEE Transactions on Knowledge and Data Engineering*, 11(1) pp. 95–107, 1999.
64. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. Using unbalanced Trees for indexing multidimensional objects. *Knowledge and Information Systems Journal*, 1(3), pp. 157–192, 1999.
65. Charu C. Aggarwal, Ravindra K. Ahuja, Jianxiu Hao, and James B. Orlin, Diagnosing Infeasibilities in Network Flow Problems, *Mathematical Programming*, 81, pp. 263–280, 1998, *Elsevier Science*, Amsterdam, Netherlands.
66. Charu C. Aggarwal, Joel L. Wolf, and Philip S. Yu. Optimization Issues in Multimedia Systems. *International Journal of Intelligent Systems on Multimedia Computing Systems*, December 1998. John Wiley and Sons.
67. Charu C. Aggarwal, James B. Orlin, and Ray P. Tai, Optimized Crossover for the Independent Set Problem. *Operations Research*, 45(2), pp. 226–234, 1997.
68. Charu C. Aggarwal, Nitesh Jain, and Phalguni Gupta, An efficient selection algorithm on the pyramid, *Information Processing Letters*, 53(1) pp. 37–47, 1995, *Elsevier Science*, Amsterdam, Netherlands.

Refereed Conferences with Published Proceedings

1. Karthik Subbian, Charu C. Aggarwal, and Kshiteesh Hegde. Recommendations for Streaming Data. *ACM CIKM Conference*, 2016.
2. Suhang Wang, Jiliang Tang, Charu C. Aggarwal, and Huan Liu. Linked Document Embedding for Classification. *ACM CIKM Conference*, 2016.
3. Arijit Khan and Charu C. Aggarwal. Query-Friendly Compression of Graph Streams. *ASONAM Conference*, 2016.
4. Jiawei Zhang, Qianyi Zhang, Lifang Hez, Charu C. Aggarwal, and Philip S. Yu. Trust Hole Identification in Signed Networks. *European Conference on Machine Learning/PKDD*, 2016.
5. Peixiang Zhao, Charu C. Aggarwal, and Gewen He. Link Prediction in Graph Streams. *IEEE International Conference on Data Engineering (ICDE)*, 2016.
6. Charu C. Aggarwal, Peixiang Zhao, and Gewen He. Edge Classification in Networks. *IEEE International Conference on Data Engineering (ICDE)*, 2016.
7. Renjun Hu, Charu C. Aggarwal, Shuai Ma, and Jinpeng Huai. An Embedding Approach to Anomaly Detection. *IEEE International Conference on Data Engineering (ICDE)*, 2016.
8. Ahsanul Haque, Latifur Khan, Michael Baron, Bhavani Thuraisingham, and Charu C. Aggarwal. Efficient Handling of Concept Drift and Concept Evolution over Stream Data. *IEEE International Conference on Data Engineering (ICDE)*, 2016.
9. Saket Sathe and Charu C. Aggarwal. LODS: Local Density meets Spectral Outlier Detection. *SIAM Conference on Data Mining*, 2016.
10. Jiliang Tang, Charu C. Aggarwal and Huan Liu. Node Classification in Signed Social Networks. *SIAM Conference on Data Mining*, 2016.
11. X. Liu, Charu C. Aggarwal, Yu-Feng Li, Xiangnan Kong, Xinyuan Sun, and Saket Sathe. Kernelized Matrix Factorization for Collaborative Filtering. *SIAM Conference on Data Mining*, 2016.
12. Shuochao Yao, Md. Tanvir Al Amin, Lu Su, Shaohan Hu, Shen Li, Shiguang Wang, Yiran Zhao, Tarek F. Abdelzaher, Lance M. Kaplan, Charu C. Aggarwal, and Aylin Yener. Recursive Ground Truth Estimator for Social Data Streams. *IPSN Conference*, pp. 1–12, 2016.
13. Jiliang Tang, Charu C. Aggarwal, and Huan Liu. Recommendations in Signed Social Networks. *World Wide Web (WWW) Conference*, pp. 31–40, 2016.
14. Karthik Subbian, Charu C. Aggarwal, and Jaideep Srivastava. Querying and Tracking Influencers in Social Streams. *Web Search and Data Mining Conference (WSDM)*, pp. 493–502, 2016.
15. Zhaoming Wu, Charu C. Aggarwal, and Jimeng Sun. The Troll-Trust Model for Ranking in Signed Networks. *Web Search and Data Mining Conference (WSDM)*, pp. 447–456, 2016.
16. Liang Duan, Charu C. Aggarwal, Shuai Ma, Renjun Hu, and Jinpeng Huai. Scaling up Link Prediction with Ensembles. *Web Search and Data Mining Conference (WSDM)*, pp. 367–376, 2016.
17. Charu C. Aggarwal and Philip S. Yu. On Historical Diagnosis of Sensor Streams. *International Conference on Data Engineering (ICDE)*, pp. 185–194, 2015.
18. Wei Feng, Chao Zhang, Wei Zhang, Jiawei Han, Jianyong Wang, Charu C. Aggarwal, and Jianbin Huang. STREAMCUBE: Hierarchical spatio-temporal hashtag clustering for event exploration over the Twitter stream. *International Conference on Data Engineering (ICDE)*, pp. 1561–1572, 2015.
19. Shiyu Chang, Wei Han, Jiliang Tang, Guo-Jun Qi, Charu C. Aggarwal, and Thomas S. Huang. Heterogeneous Network Embedding via Deep Architectures. *ACM KDD Conference*, pp. 119–128, 2015.

20. Guo-Jun Qi, Charu C. Aggarwal, Deepak S. Turaga, Daby M. Sow, and Phil Anno. State-Driven Dynamic Sensor Selection and Prediction with State-Stacked Sparseness. *ACM KDD Conference*, pp. 945–954, 2015.
21. Jialu Liu, Chi Wang, Jing Gao, Quanquan Gu, Charu C. Aggarwal, Lance M. Kaplan, and Jiawei Han. GIN: A Clustering Model for Capturing Dual Heterogeneity in Networked Data. *SIAM Conference on Data Mining (SDM)*, pp. 388–396, 2015.
22. Karthik Subbian, Charu C. Aggarwal, Jaideep Srivastava, and Vipin Kumar. Rare Class Detection in Networks. *SIAM Conference on Data Mining (SDM)*, pp. 406–414, 2015.
23. Min-Hsuan Tsai, Charu C. Aggarwal, and Thomas S. Huang. Towards Classification of Social Streams. *SIAM Conference on Data Mining (SDM)*, pp. 649–657, 2015.
24. Hessam Zakerzadeh, Charu C. Aggarwal, and Ken Barker. Privacy-preserving Big Data Publishing. *International Conference on Scientific and Statistical Data Management (SSDBM)*, 26:1–26:11, 2015.
25. Jiliang Tang, Shiyu Chang, Charu C. Aggarwal, and Huan Liu. Negative Link Prediction in Social Media. *Web Search and Data Mining (WSDM) Conference*, pp. 87–96, 2015.
26. Jialu Liu, Charu C. Aggarwal, and Jiawei Han. On Integrating Network and Community Discovery. *Web Search and Data Mining (WSDM) Conference*, pp. 117–126, 2015.
27. Ahsanul Haque, Swarup Chandra, Latifur Khan, Charu Aggarwal: Distributed Adaptive Importance Sampling on graphical models using MapReduce. *IEEE BigData Conference*, pp. 597–602, 2014.
28. Tyler Clemons, S. M. Faisal, Shirish Tatikonda, Charu C. Aggarwal, Srinivasan Parthasarathy: One, Two, Hash! Counting Hash Tables for Flash Devices. *CODS*, pp. 1–10, 2014.
29. Siyu Gu, Chenji Pan, Hengchang Liu, Shen Li, Shaohan Hu, Lu Su, Shiguang Wang, Dong Wang, Md. Tanvir Al Amin, Ramesh Govindan, Charu C. Aggarwal, Raghu K. Ganti, Mudhakar Srivatsa, Amotz Barnoy, Peter Terlecky, and Tarek F. Abdelzaher. Data Extrapolation in Social Sensing for Disaster Response. *International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 119–126, 2014.
30. Shiyu Chang, Guo-Jun Qi, Charu C. Aggarwal, Jiayu Zhou, Meng Wang, and Thomas S. Huang. Factorized Similarity Learning in Networks. *International Conference on Data Mining (ICDM)*, pp. 60–69, 2014.
31. Shiyu Chang, Charu C. Aggarwal, and Thomas S. Huang. Learning Local Semantic Distances with Limited Supervision. *International Conference on Data Mining (ICDM)*, pp. 70–79, 2014.
32. Swarup Chandra, Justin Sahs, Latifur Khan, Bhavani M. Thuraisingham, and Charu C. Aggarwal. Stream Mining Using Statistical Relational Learning. *International Conference on Data Mining (ICDM)*, pp. 743–748, 2014.
33. Dong Wang, Md. Tanvir Al Amin, Shen Li, Tarek F. Abdelzaher, Lance M. Kaplan, Siyu Gu, Chenji Pan, Hengchang Liu, Charu C. Aggarwal, Raghu K. Ganti, Xinlei Wang, Prasant Mohapatra, Boleslaw K. Szymanski, Hieu Khac Le: Using humans as sensors: an estimation-theoretic perspective. *International Conference on Information Processing in Sensor Networks (IPSN)*, pp. 35–46, 2014.
34. Charu C. Aggarwal. The Setwise Stream Classification Problem. *ACM KDD Conference*, pp. 432–441, 2014.
35. Karthik Subbian, Chidananda Sridhar, Charu C. Aggarwal, Jaideep Srivastava: Scalable Information Flow Mining in Networks. *ECML/PKDD Conference*, pp. 130–146, 2014.
36. Min-Hsuan Tsai, Charu C. Aggarwal and Thomas Huang. Ranking in Heterogeneous Social Media, *Web Search and Data Mining Conference (WSDM)*, pp. 613–622, 2014.

37. Hessam Zakerzadeh, Charu C. Aggarwal and Ken Barker. Towards Breaking the Curse of Dimensionality for High-Dimensional Privacy, *SIAM Conference on Data Mining (SDM)*, pp. 731–739, 2014.
38. Weiren Yu, Charu C. Aggarwal, Shuai Ma, and Haixun Wang, On Anomalous Hot Spot Discovery in Graph Streams, *IEEE International Conference on Data Mining (ICDM)*, pp. 171–1276, 2013.
39. Karthik Subbian, Charu C. Aggarwal, Jaideep Srivasatava. Content-centric Flow Mining for Influence Analysis in Social Streams, *ACM Conference on Information and Knowledge Management (CIKM)*, pp. 841–846, 2013.
40. Quanquan Gu, Charu C. Aggarwal, and Jiawei Han. Unsupervised Link Selection in Networks. *AISTATS Conference*, pp. 298–306, 2013.
41. Tyler Clemons, S. M. Faisal, Shirish Tatikonda, Charu C. Aggarwal, and Srinivasan Parthasarathy. Hash in a Flash: Hash Tables for Flash Devices. *IEEE BigData Conference*, pp. 7–14, 2013.
42. Dong Wang, Tarek F. Abdelzaher, Lance M. Kaplan, and Charu C. Aggarwal. Recursive Fact-Finding: A Streaming Approach to Truth Estimation in Crowdsourcing Applications. *International Conference on Distributed Computing Systems (ICDCS)*, pp. 530–539, 2013.
43. Guo-Jun Qi, Charu C. Aggarwal, and Thomas S. Huang. Link prediction across networks by biased cross-network sampling. *International Conference on Data Engineering (ICDE)*, pp. 793–804, 2013.
44. Quanquan Gu, Charu C. Aggarwal, Jialu Liu, and Jiawei Han. Selective sampling on graphs for classification. *ACM KDD Conference*, pp. 131–139, 2013.
45. Charu C. Aggarwal, Jaideep Srivastava, Karthik Subbian, and Philip S. Yu. Community Detection with Prior Knowledge. *SIAM Conference on Data Mining (SDM)*, pp. 405–413, 2013.
46. Guo-Jun Qi, Charu C. Aggarwal, and Thomas S. Huang. Online community detection in social sensing. *Web Search and Data Mining (WSDM) Conference*, pp. 617–626, 2013.
47. Guo-Jun Qi, Charu C. Aggarwal, Jiawei Han, and Thomas S. Huang. Mining collective intelligence in diverse groups. *Worldwide Web (WWW) Conference*, pp. 1041–1052, 2013.
48. Arijit Khan, Yinghui Wu, Charu C. Aggarwal, and Xifeng Yan. NeMa: Fast Graph Search with Label Similarity, *Very Large Databases Conference (VLDB)*, pp. 181–192, 2013.
49. Guojun Qi, Charu Aggarwal, and Thomas Huang. Link Prediction across Networks by Cross-Network Biased Sampling. *International Conference on Data Engineering*, pp. 793–804, 2013.
50. Lin Liu, Ruoming Jin, Charu C. Aggarwal, and Yelong Shen. Reliable Clustering of Uncertain Graphs, *ICDM Conference*, pp. 459–468, 2012.
51. Tahseen M. Al-Khateeb, Mohammad M. Masud, Latifur Khan, Charu C. Aggarwal, Jiawei Han, and Bhavani Thuraisingham. Stream Classification with Recurring and Novel Class Detection using Class-Based Ensemble. *International Conference on Data Mining (ICDM)*, 31–40, 2012.
52. Charu C. Aggarwal, Wangqun Lin, Philip S. Yu. Searching by Corpus with Fingerprints, *International Conference on Extending Database Technology (EDBT)*, pp. 348–359, 2012.
53. Dong Wang, Lance M. Kaplan, Tarek F. Abdelzaher, and Charu C. Aggarwal. On Scalability and Robustness Limitations of Real and Asymptotic Confidence Bounds in Social Sensing. *SECON Conference*, pp. 506–514, 2012.
54. Charu C. Aggarwal, Karthik Subbian. Event Detection in Social Streams, *SIAM Conference on Data Mining (SDM)*, pp. 624–635, 2012.
55. Charu C. Aggarwal, Yan Xie, Philip S. Yu. On Dynamic Link Inference in Heterogeneous Networks, *SIAM Conference on Data Mining (SDM)*, pp. 415–426, 2012.
56. Charu C. Aggarwal, Shuyang Lin, Philip S. Yu. On Influential Node Discovery in Dynamic Social Networks, *SIAM Conference on Data Mining (SDM)*, pp. 636–647, 2012.

57. Charu C. Aggarwal. The Multi-Set Stream Clustering Problem, *SIAM Conference on Data Mining (SDM)*, pp. 59–69, 2012.
58. GuoJun Qi, Charu C. Aggarwal, Thomas S. Huang. Transfer Learning of Distance Metrics by Cross-Domain Metric Sampling across Heterogeneous Domains, *SIAM Conference on Data Mining (SDM)*, pp. 528–539, 2012.
59. GuoJun Qi, Charu C. Aggarwal, Thomas S. Huang. Community Detection with Edge Content in Social Media Networks, *IEEE International Conference on Data Engineering (ICDE)*, pp. 534–545, 2012.
60. Charu C. Aggarwal, Yuchen Zhao, Philip S. Yu. On Text Clustering with Side Information, *IEEE International Conference on Data Engineering (ICDE)*, pp. 894–904, 2012.
61. GuoJun Qi, Charu C. Aggarwal, Thomas S. Huang. On clustering heterogeneous social media objects with outlier links. *Web Search and Data Mining (WSDM) Conference*, pp. 553–562, 2012.
62. Yizhou Sun, Jiawei Han, Charu C. Aggarwal, Nitesh V. Chawla. When will it happen?: relationship prediction in heterogeneous information networks. *Web Search and Data Mining (WSDM) Conference*, pp. 663–672, 2012.
63. Yizhou Sun, Rick Barber, Manish Gupta, Charu C. Aggarwal, Jiawei Han. Co-author Relationship Prediction in Heterogeneous Bibliographic Networks. *ASONAM Conference*, pp. 121–128, 2011.
64. Manish Gupta, Charu C. Aggarwal, Jiawei Han, Yizhou Sun. Evolutionary Clustering and Analysis of Bibliographic Networks. *ASONAM Conference*, pp. 63–70, 2011. **(Best Paper Award)**
65. Guo-Jun Qi, Charu C. Aggarwal, Yong Rui, Qi Tian, Shiyu Chang, Thomas S. Huang. Towards cross-category knowledge propagation for learning visual concepts. *Computer Vision and Pattern Recognition (CVPR) Conference*, pp. 897–904, 2011.
66. Charu C. Aggarwal, Amotz Bar-Noy, Simon Shamoun. On sensor selection in linked information networks. *International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2011.
67. Dong Wang, Hossein Ahmadi, Tarek F. Abdelzaher, Harshvardhan Chenji, Radu Stoleru, Charu C. Aggarwal. Optimizing quality-of-information in cost-sensitive sensor data fusion. *International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2011.
68. Charu C. Aggarwal, Haixun Wang. On dimensionality reduction of massive graphs for indexing and retrieval. *International Conference on Data Engineering (ICDE)*, pp. 1091–1102, 2011.
69. Charu C. Aggarwal, Yuchen Zhao, Philip S. Yu. Outlier detection in graph streams. *ICDE Conference*, pp. 399–409, 2011.
70. Charu C. Aggarwal, Yao Li, Philip S. Yu. On the Hardness of Graph Anonymization. *International Conference on Data Mining (ICDM)*, pp. 1002–1007, 2011.
71. Mohammad M. Masud, Tahseen Al-Khateeb, Latifur Khan, Charu C. Aggarwal, Jing Gao, Jiawei Han, Bhavani M. Thuraisingham. Detecting Recurring and Novel Classes in Concept-Drifting Data Streams. *International Conference on Data Mining (ICDM)*, pp. 1176–1181, 2011.
72. Charu C. Aggarwal, Yan Xie, Philip S. Yu. On dynamic data-driven selection of sensor streams. *ACM KDD Conference*, pp. 1226–1234, 2011.
73. Ruoming Jin, Lin Liu, Charu C. Aggarwal. Discovering highly reliable subgraphs in uncertain graphs. *ACM KDD Conference*, pp. 992–1000, 2011.
74. Charu C. Aggarwal, Nan Li. On Node Classification in Dynamic Content-based Networks. *SIAM Conference on Data Mining (SDM)*, pp. 355–366, 2011.
75. Charu C. Aggarwal, Yan Xie, Philip S. Yu. Towards Community Detection in Locally Heterogeneous Networks. *SIAM Conference on Data Mining (SDM)*, pp. 391–402.
76. Charu C. Aggarwal, Arijit Khan, Xifeng Yan. On Flow Authority Discovery in Social Networks. *SIAM Conference on Data Mining (SDM)*, pp. 522–533, 2011.

77. Charu C. Aggarwal. On Classification of Graph Streams. *SDM Conference*, pp. 652–663, 2011.
78. Chun Li, Charu C. Aggarwal, Jianyong Wang: On Anonymization of Multi-graphs. *SIAM Conference on Data Mining (SDM)*, pp. 711–722, 2011.
79. Manish Gupta, Charu C. Aggarwal, Jiawei Han. Finding Top-k Shortest Path Distance Changes in an Evolutionary Network. *Symposium on Spatial and Temporal Data (SSTD)*, pp. 130–148, 2011.
80. Guojun Qi, Charu C. Aggarwal, Thomas S. Huang. Towards semantic knowledge propagation from text corpus to web images. *World Wide Web (WWW) Conference*, pp. 297–306, 2011.
81. Yuchen Zhao, Charu C. Aggarwal, Philip S. Yu. On wavelet decomposition of uncertain time series data sets. *ACM CIKM Conference*, pp. 129–138, 2010.
82. Mohammad M. Masud, Qing Chen, Latifur Khan, Charu C. Aggarwal, Jing Gao, Jiawei Han, Bhavani M. Thuraisingham. Addressing Concept-Evolution in Concept-Drifting Data Streams. *International Conference on Data Mining (ICDM)*, pp. 929–934, 2010.
83. Charu C. Aggarwal. On Multidimensional Sharpening of Uncertain Data. *SIAM Conference on Data Mining (SDM)*, pp. 373–384, 2010.
84. Charu C. Aggarwal, Yuchen Zhao, Philip S. Yu. On Clustering Graph Streams. *SIAM Conference on Data Mining (SDM)*, pp. 478–489, 2010.
85. Charu C. Aggarwal. The Generalized Dimensionality Reduction Problem. *SIAM Conference on Data Mining (SDM)*, pp. 607–618, 2010.
86. Charu C. Aggarwal, Philip S. Yu. On Classification of High-Cardinality Data Streams. *SIAM Conference on Data Mining (SDM)*, pp. 802–813, 2010.
87. Charu C. Aggarwal, Peixiang Zhao. Graphical models for text: a new paradigm for text representation and processing. *ACM SIGIR Conference*, pp. 899–900, 2010.
88. Charu C. Aggarwal, Philip S. Yu. Online Auctions: There Can Be Only One. *IEEE International Conference on Electronic Commerce (CEC)*, pp. 176–181, 2009.
89. Charu C. Aggarwal, Yan Li, Jianyong Wang, Jing Wang. Frequent pattern mining with uncertain data. *ACM KDD Conference*, pp. 29–38, 2009.
90. Charu C. Aggarwal, A Framework for Clustering Massive-Domain Data Streams, *IEEE International Conference on Data Engineering (ICDE)*, 2009.
91. Charu C. Aggarwal, On High-Dimensional Projected Clustering of Uncertain Data Streams, *IEEE International Conference on Data Engineering (ICDE)*, 2009.
92. Dina Thomas, Rajesh Bordawekar, Charu Aggarwal, Philip S. Yu, On Efficient Query Processing of Stream Counts on the Cell Processor, *IEEE International Conference on Data Engineering (ICDE)*, 2009.
93. Charu C. Aggarwal, On Segment-Based Stream Modeling and its Applications, *SIAM Conference on Data Mining (SDM)*, pp. 721–732, 2009.
94. Charu C. Aggarwal, Philip S. Yu. On Indexing High Dimensional Data with Uncertainty, *SIAM Conference on Data Mining (SDM)*, 2008.
95. Charu C. Aggarwal, Philip S. Yu. Outlier Detection with Uncertain Data. *SIAM Conference on Data Mining (SDM)*, 2008.
96. Charu C. Aggarwal, Philip S. Yu. On High Dimensional Indexing of Uncertain Data. *ICDE Conference*, 2008.
97. Charu C. Aggarwal. On Unifying Privacy and Uncertain Data Models. *IEEE International Conference on Data Engineering (ICDE)*, 2008.

98. Charu C. Aggarwal, Philip S. Yu. LOCUST: An Online Analytical Processing Framework for High Dimensional Classification of Data Streams. *IEEE International Conference on Data Engineering (ICDE)*, 2008.
99. Charu C. Aggarwal, Philip S. Yu. A Framework for Clustering Uncertain Data Streams. *IEEE International Conference on Data Engineering (ICDE)*, 2008.
100. Kun-Lung Wu, Philip S. Yu, Bugra Gedik, Kirsten Hildrum, Charu C. Aggarwal, Eric Bouillet, Wei Fan, David George, Xiaohui Gu, Gang Luo, Haixun Wang: Challenges and Experience in Prototyping a Multi-Modal Stream Analytic and Monitoring Application on System S. *Very Large Databases (VLDB) Conference*, pp. 1185-1196, 2007.
101. Charu C. Aggarwal, Philip S. Yu. On String Classification in Data Streams. *ACM KDD Conference*, pp. 36-45, 2007.
102. Charu C. Aggarwal, Na Ta, Jianyong Wang, Jianhua Feng, Mohammed Zaki. XProj: A Framework for Projected Structural Clustering of XML Documents. *ACM KDD Conference*, pp. 46-55, 2007.
103. Charu C. Aggarwal. A Framework for Classification and Segmentation of Massive Audio Data Streams. *ACM KDD Conference*, pp. 1013-1017, 2007.
104. Charu C. Aggarwal. On Point Sampling versus Space Sampling for Dimensionality Reduction. *SIAM Conference on Data Mining (SDM)*, 2007.
105. Charu C. Aggarwal, Philip S. Yu. On Privacy-Preservation of Text and Sparse Binary Data with Sketches. *SIAM Conference on Data Mining (SDM)*, 2007.
106. Charu C. Aggarwal, Philip S. Yu. On Anonymization of String Data. *SIAM Conference On Data Mining (SDM)*, 2007.
107. Jianyong Wang, Yuzhou Zhang, Lizhu Zhou, George Karypis, Charu C. Aggarwal. Discriminating subsequence discovery for sequence clustering. *SIAM Conference on Data Mining (SDM)*, 2007.
108. Charu C. Aggarwal. On Randomization, Public Information, and the Curse of Dimensionality. *IEEE International Conference on Data Engineering*, pp. 136-145, 2007.
109. Charu C. Aggarwal. On Density Based Transforms for Uncertain Data Mining. *IEEE International Conference on Data Engineering (ICDE)*, pp. 866-875, 2007.
110. Charu C. Aggarwal. On Biased Reservoir Sampling in the Presence of Stream Evolution. *Very Large Databases (VLDB) Conference*, pp. 607-618, 2006.
111. Charu C. Aggarwal, Jian Pei, Bo Zhang. On Privacy Preservation against Adversarial Data Mining. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 510-516, 2006.
112. Charu C. Aggarwal. On Futuristic Query Processing in Data Streams. *International Conference on Extending Database Technology (EDBT)*, pp. 41-58, 2006.
113. Charu C. Aggarwal, Chen Chen, Jiawei Han. On the Inverse Classification Problem and its Applications. *IEEE International Conference on Data Engineering (ICDE)*, 2006.
114. Charu C. Aggarwal. A Framework for Local Supervised Dimensionality Reduction of High Dimensional Data, *SIAM Conference on Data Mining (SDM)*, pp. 360-371, 2006.
115. Charu C. Aggarwal. Representation is Everything: Towards Efficient and Adaptable Similarity Measures for Biological Data , *SIAM Conference on Data Mining (SDM)*, pp. 210-221, 2006.
116. Charu C. Aggarwal, Philip S. Yu. A Framework for Clustering Massive Text and Categorical Data Streams, *SIAM Conference on Data Mining (SDM)*, pp. 479-483, 2006.
117. Charu C. Aggarwal. On k -anonymity and the curse of dimensionality. *Very Large Databases (VLDB) Conference*, pp. 901-909, 2005.
118. Charu C. Aggarwal. Towards Exploratory Test Instance Centered Algorithms for High Dimensional Classification. *Knowledge Discovery and Data Mining (KDD) Conference*, pp. 526-531, 2005.

119. Charu C. Aggarwal, Philip S. Yu. Online Analysis of Community Evolution in Data Streams. *SIAM Conference on Data Mining*, pp. 56–67, 2005.
120. Charu C. Aggarwal, Philip S. Yu. On Variable Constraints in Privacy Preserving Data Mining Algorithms. *SIAM Conference on Data Mining (SDM)*, pp. 115–125, 2005.
121. Charu C. Aggarwal. On Abnormality Detection in Spuriously Populated Data Streams. *SIAM Conference on Data Mining (SDM)*, pp. 80–91, 2005.
122. Charu C. Aggarwal, Philip S. Yu. A Condensation Approach to Privacy Preserving Data Mining. *Conference on Extending Database Technology (EDBT)*, pp. 183–199, 2004.
123. Charu C. Aggarwal, Jiawei Han, Jianyong Wang, Philip S. Yu. On Demand Classification of Data Streams. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 503–508, 2004.
124. Charu C. Aggarwal, Jiawei Han, Jianyong Wang, Philip S. Yu. A Framework For Projected Clustering of High Dimensional Data Streams. *Very Large Databases (VLDB) Conference*, pp. 852–863, 2004.
125. Charu C. Aggarwal. A Framework for Diagnosing Changes in Evolving Data Streams. *ACM SIGMOD Conference on Management of Data*, pp. 575–586, 2003.
126. Charu C. Aggarwal, Jiawei Han, Jianyong Wang, Philip Yu. A Framework for Clustering Evolving Data Streams. *Very Large Databases (VLDB) Conference*, pp. 81–92, 2003.
127. Charu C. Aggarwal. Towards Systematic Design of Distance Functions for Data Mining Applications. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 9–18, 2003.
128. M. J. Zaki, Charu C. Aggarwal. XRules: An Effective Structural Classifier for XML Data. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 316–325, 2003.
129. Charu C. Aggarwal, Dakshi Agrawal. On Nearest Neighbor Indexing of Nonlinear Trajectories. *ACM Principles on Database Systems (PODS Conference)*, pp. 252–259, 2003.
130. Charu C. Aggarwal, Philip S. Yu. An Automated System for Web Portal Personalization. *Very Large Databases (VLDB) Conference*, pp. 1031–1040, 2002.
131. Charu C. Aggarwal. On Effective Classification of Strings with Wavelets. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 163–172, 2002.
132. Charu C. Aggarwal. Collaborative Crawling: Mining User Experiences for Topical Resource Discovery. *ACM Knowledge Discovery and Data Mining (KDD Conference)*, pp. 423–428, 2002.
133. Charu C. Aggarwal. Towards Meaningful High Dimensional Nearest Neighbor Search by Human-Computer Interaction. *IEEE International Conference On Data Engineering (ICDE)*, pp. 593–604, 2002.
134. Charu C. Aggarwal. An Intuitive Framework for Understanding Changes in Evolving Data Streams. *IEEE International Conference on Data Engineering*, 2002.
135. Charu C. Aggarwal, Philip S. Yu. A System for Personalization of Web News Feeds. *IEEE International Conference on Data Engineering (ICDE)*, 2002.
136. Charu C. Aggarwal. Hierarchical Subspace Sampling: A Unified Framework for High Dimensional Data Reduction, Selectivity Estimation and Nearest Neighbor Search. *ACM SIGMOD Conference on Management of Data*, pp. 452–463, 2002.
137. Charu C. Aggarwal. On the Effects of Dimensionality Reduction on High Dimensional Similarity Search. *ACM Principles on Databases Systems (PODS Conference)*, pp. 256–266, 2001.
138. Dakshi Agrawal, Charu C. Aggarwal. On the Design and Quantification of Privacy Preserving Data Mining Algorithms. *ACM Principles on Databases Systems (PODS Conference)*, pp. 247–255, 2001.

139. Charu C. Aggarwal. A Human-Computer Cooperative System for Effective High Dimensional Clustering. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 221–226, 2001.
140. Charu C. Aggarwal, Srinivasan Parthasarathy. Mining Massively Incomplete Data Sets by Conceptual Reconstruction. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 227–232, 2001.
141. Charu C. Aggarwal. Towards Long Pattern Generation in Dense Databases. *ACM SIGKDD Explorations*, 3(1), pp. 20–26, 2001.
142. Charu C. Aggarwal, Philip S. Yu. On Effective Conceptual Indexing and Similarity Search in Text Data. *IEEE International Conference on Data Mining (ICDM Conference)*, pp. 3–10, 2001.
143. Sang-Wook Kim, Charu C. Aggarwal, Philip S. Yu. Effective Nearest Neighbor Indexing with the Euclidean Metric. *ACM CIKM Conference*, pp. 9–16, 2001.
144. Charu C. Aggarwal, Philip S. Yu. Outlier Detection for High Dimensional Data. *ACM SIGMOD Conference on Management of Data*, pp. 37–46, 2001.
145. Charu C. Aggarwal, Fatima Al-Garawi, Philip S. Yu. Intelligent Crawling on the World Wide Web with arbitrary predicates. *World Wide Web (WWW) Conference*, pp. 96–105, 2001.
146. Charu C. Aggarwal, Alexander Hinneburg, Daniel A. Keim. On the surprising behavior of distance metrics in high dimensional space. *International Conference on Database Theory (ICDT Conference)*, pp. 420–434, 2001.
147. Alexander Hinneburg, Charu C. Aggarwal, Daniel A. Keim. What is the nearest neighbor in high dimensional space? *Very Large Databases (VLDB) Conference*, pp. 506–515, 2000.
148. Charu C. Aggarwal, Philip S. Yu. The IGrid Index: Reversing the dimensionality curse for similarity indexing in high dimensional space. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 119–129, 2000.
149. Ramesh C. Agarwal, Charu C. Aggarwal, V.V.V. Prasad. Depth First Generation of Long Patterns. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 108–118, 2000.
150. Charu C. Aggarwal, Philip S. Yu. Finding Generalized Projected Clusters in High dimensional spaces. *ACM SIGMOD Conference on Management of Data*, pp. 70–81, 2000.
151. Charu C. Aggarwal, Joel L. Wolf, Kun-Lung Wu, Philip S. Yu. Horting Hatches an Egg: A New Graph Theoretic Approach to Collaborative Filtering. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 201–212, 1999.
152. Charu C. Aggarwal, Stephen C. Gates, Philip S. Yu. On the merits of building categorization systems by supervised clustering. *ACM Knowledge Discovery and Data Mining (KDD) Conference*, pp. 352–356, 1999.
153. Charu C. Aggarwal, Cecilia Procopiuc, Joel L. Wolf, Philip S. Yu, Jong Soo Park. Fast Algorithms for Projected Clustering. *ACM SIGMOD Conference on Management of Data*, pp. 61–72, 1999.
154. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. A new method for similarity indexing of market basket data. *ACM SIGMOD Conference on Management of Data*, pp. 407–418, 1999.
155. Charu C. Aggarwal, Philip S. Yu. Online Algorithms for Finding Profile Association Rules. *ACM CIKM Conference*, pp. 86–95, 1998.
156. Charu C. Aggarwal, Zheng Sun, Philip S. Yu. Online Generation of Profile Association Rules. *Knowledge Discovery and Data Mining (KDD) Conference*, pp. 129–133, 1998.
157. Charu C. Aggarwal, Joel L. Wolf and Philip S. Yu, A Framework for the Optimizing of WWW Advertising, *Trends in Distributed Systems for Electronic Commerce, Proceedings in Lecture Notes in Computer Science*, Vol. 1402, pp. 1–10, 1998.
158. Charu C. Aggarwal and Philip S. Yu. Online Generation of Association Rules. *IEEE International Conference on Data Engineering*, pp. 402–411, 1998.

159. Charu C. Aggarwal, and Philip S. Yu, Mining large itemsets for association rules. *IEEE Data Engineering Bulletin*, 21(1), pp. 23–31, 1998.
160. Charu C. Aggarwal, and Philip S. Yu, A new framework for itemset generation. *ACM Symposium on Principles of Database Systems (PODS Conference)*, pp. 18–24, 1998.
161. Charu C. Aggarwal and Philip S. Yu. On disk caching of web objects in proxy servers. *International Conference on Information and Knowledge Management*, pp. 238–245, 1997.
162. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu, and Marina Epelman. The S-Tree: An efficient index for multi-dimensional objects. *Symposium on Spatial Databases*, pp. 350–376, 1997.
163. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu, On optimal piggyback merging policies in video-on-demand systems. *ACM SIGMETRICS Conference*, pp. 200–209, 1996.
164. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu, On optimal batching policies in video-on-demand storage systems. *IEEE International Conference on Multimedia Computing and Systems*, pp. 253–258, 1996.
165. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu, A permutation based pyramid broadcasting scheme for video-on-demand storage systems. *IEEE International Conference on Multimedia Computing and Systems*, pp. 118–126, 1996.

Funding Proposals

Co-PI on *Information Networks Academic Research Center*, Network Science Proposal, Army Research Labs, Joint proposal with UIUC, UCSB, and CUNY, 30 million dollars (2009).

Book Chapters

1. Charu C. Aggarwal, Deepak Turaga. Mining Data Streams: Systems and Algorithms, *Machine Learning and Knowledge Discovery for Engineering Systems Health Management*, ed. Ashok Srivastava, Jiawei Han, CRC Press, 2012.
2. Charu C. Aggarwal, ChengXiang Zhai. An Introduction to Text Mining. *Mining Text Data*, Springer, pp. 1–10, 2012.
3. Charu C. Aggarwal, ChengXiang Zhai. A Survey of Text Classification Algorithms. *Mining Text Data*, Springer, pp. 163–222, 2012.
4. Charu C. Aggarwal. Mining Text Streams, *Mining Text Data*, Springer, pp. 163–222, 2012.
5. Charu C. Aggarwal, ChengXiang Zhai. A Survey of Text Clustering Algorithms. *Mining Text Data*, Springer, pp. 77–128, 2012.
6. Charu C. Aggarwal. An Introduction to Social Network Data Analytics. *Social Network Data Analytics*, Springer, pp. 1–15, 2011.
7. Charu C. Aggarwal, Haixun Wang. Text Mining in Social Networks. *Social Network Data Analytics*, Springer, pp. 353–378, 2011.
8. Charu C. Aggarwal, Tarek F. Abdelzaher. Integrating Sensors and Social Networks. *Social Network Data Analytics*, Springer, pp. 379–412, 2011.
9. Charu C. Aggarwal, Haixun Wang. An Introduction to Graph Data, *Managing and Mining Graph Data*, Springer, 2010.
10. Charu C. Aggarwal, Haixun Wang. Graph Data Management and Mining: A Survey of Algorithms and Applications. *Managing and Mining Graph Data*, Springer, pp. 13–68, 2010.
11. Haixun Wang, Charu C. Aggarwal. A Survey of Algorithms for Keyword Search on Graph Data. *Managing and Mining Graph Data*, Springer, pp. 249–273, 2010.
12. Charu C. Aggarwal, Haixun Wang. A Survey of Clustering Algorithms for Graph Data. *Managing and Mining Graph Data*, Springer, pp. 275–301, 2010.

13. Victor E. Lee, Ning Ruan, Ruoming Jin, Charu C. Aggarwal. A Survey of Algorithms for Dense Subgraph Discovery. *Managing and Mining Graph Data*, Springer, pp. 303–336, 2010.
14. Charu C. Aggarwal. An Introduction to Uncertain Data Mining, *Managing and Mining Uncertain Data*, Springer, 2009.
15. Charu C. Aggarwal. On Clustering Algorithms for Uncertain Data, *Managing and Mining Uncertain Data*, Springer, pp. 389–406, 2009.
16. Charu C. Aggarwal. On Applications of Density Transforms for Uncertain Data Mining, *Managing and Mining Uncertain Data*, Springer, pp. 407–424, 2009.
17. Charu C. Aggarwal, Yan Li, Jianyong Wang, Jing Wang. Frequent Pattern Mining Algorithms with Uncertain Data, *Managing and Mining Uncertain Data*, Springer, pp. 427–460, 2009.
18. Charu C. Aggarwal, Stephen C. Gates, Philip S. Yu. On Supervised Clustering for Creating Categorization Segmentations. In *Constrained Clustering: Advances in Theory, Algorithms and Applications*. Ed. Sugato Basu, Ian Davidson, Kiri Wagstaff, Chapman and Hall, 2008.
19. Charu C. Aggarwal, Philip S. Yu. An Introduction to Privacy-Preserving Data Mining. *Privacy-Preserving Data Mining: Models and Algorithms*. Ed. Charu Aggarwal, Philip S. Yu, Springer Inc., 2008.
20. Charu C. Aggarwal, Philip S. Yu. A General Survey of Privacy-Preserving Data Mining Models and Algorithms. *Privacy-Preserving Data Mining: Models and Algorithms*. Ed. Charu C. Aggarwal, Philip S. Yu, Springer, 2008.
21. Charu C. Aggarwal, Philip S. Yu. A survey of randomization algorithms for privacy-preserving data mining. *Privacy-Preserving Data Mining: Models and Algorithms*. Ed. Charu C. Aggarwal, Philip S. Yu, Springer, 2008.
22. Charu C. Aggarwal. On the Curse of Dimensionality for Privacy-Preserving Data Mining. *Privacy-Preserving Data Mining: Models and Algorithms*. Ed. Charu C. Aggarwal, Philip S. Yu, Springer, 2008.
23. Charu C. Aggarwal, Philip S. Yu. Privacy-Preserving Data Mining: A Survey. *Handbook on Database Security*, ed. Michael Gertz, Sushil Jajodia, Springer, 2008.
24. Charu C. Aggarwal. An Introduction to Data Streams. *Data Streams: Models and Algorithms*. Ed. Charu Aggarwal, Springer Inc., 2007.
25. Charu C. Aggarwal, Philip S. Yu. A Survey of Synopsis Construction in Data Streams. *Data Streams: Models and Algorithms*. Ed. Charu Aggarwal, Springer Inc., 2007.
26. Charu C. Aggarwal. A Survey of Change Diagnosis Methods in Data Streams. *Data Streams: Models and Algorithms*, Ed. Charu Aggarwal, Springer Inc., 2007.
27. Charu C. Aggarwal, Jiawei Han, Jianyong Wang, Philip Yu. On Clustering Data Streams: A Summarization Paradigm. *Data Streams: Models and Algorithms*. Ed. Charu Aggarwal, Springer Inc., 2007.
28. Charu C. Aggarwal, Philip S. Yu. On Clustering Techniques for Change Diagnosis in Data Streams. *Lecture Notes in Computer Science*, vol. 4198, 2006.
29. Charu C. Aggarwal. On Learning Strategies for Topic Specific Web Crawling. *Next Generation Data Mining Applications*, Ed. J. Zurada, M. Kantardzic, John Wiley Inc., 2005.
30. Charu C. Aggarwal, Philip S. Yu. Data Mining Techniques for Associations, Clustering, and Classification, *Lecture Notes in Computer Science*, Vol. 1574, 1999.

Conference Tutorials

1. Manish Gupta, Jing Gao, Charu Aggarwal, Jiawei Han. Temporal Outlier Detection, *SIAM Conference on Data Mining (SDM)*, 2013.

2. Manish Gupta, Jing Gao, Charu Aggarwal, Jiawei Han. Outlier Detection in Graph Data, *ASONAM Conference*, 2013.
3. Manish Gupta, Jing Gao, Charu Aggarwal, Jiawei Han. Outlier Detection in Temporal Data, *Conference on Information and Knowledge Management (CIKM)*, 2013.

Patents

80 patents have been filed. Of these, 64 have already been granted by the US patent office. The remaining are pending processing at the patent office. I have also been granted several invention plateaus by IBM for my patents. I have thrice been designated **Master Inventor** by IBM Research for the commercial value of my patents. About 20% of these patents have also been filed or granted internationally, depending upon scope. A list of the granted patents is below:

1. Charu C. Aggarwal. System and Method for Finding Important Nodes in a Network, US Patent 8,553,587, 2013.
2. Charu C. Aggarwal, Peixiang Zhao Query optimization over graph data streams, US Patent 392,398, 2013.
3. Charu C. Aggarwal. Graphical Models for Representing Text Documents for Computer Analysis, US Patent 8,375,061, 2013.
4. Charu C. Aggarwal, R. Bordawekar. Identifying Communities in an Information Network, US Patent 396,855, 2013.
5. Charu C. Aggarwal, Rajesh Bordawekar, Dina Thomas, and Philip Yu. System and method for analyzing streams and counting stream items on multi-core processors, US Patent 8,321,579, 2012.
6. Charu C. Aggarwal and Philip S. Yu. System and method for classifying data streams with very large cardinality, US Patent 8,311,959, 2012.
7. Charu C. Aggarwal. Method and Apparatus for Monitoring Abnormalities in a Data Stream, US Patent 8,165,975, 2012.
8. Charu C. Aggarwal and Philip S. Yu. System and Method for Resource Adaptive Classification of Data Streams, US Patent 8,165,979, 2012.
9. Charu C. Aggarwal, Philip S. Yu. Method, apparatuses, and computer program products for classifying uncertain data, US Patent 8,086,550, 2011.
10. Charu C. Aggarwal, Philip S. Yu. Method and Apparatus for Intelligent Crawling on the World Wide Web, US Patent 8,060,816, 2011.
11. Charu C. Aggarwal, Philip S. Yu, System and Method for Resource Adaptive Classification of Data Streams, US Patent 8,051,021, 2011.
12. Charu C. Aggarwal, Philip S. Yu System and Method for Condensation-based Privacy in Strings, US Patent 8,010,541, 2011.
13. Charu C. Aggarwal, Philip S. Yu. Method and Apparatus for Aggregation in Uncertain Data, US Patent 8,005,839, 2011.
14. Charu C. Aggarwal, Philip S. Yu. Method and Apparatus for Data Stream Clustering for Abnormality Monitoring, US Patent 7,970,772, 2011.
15. Charu C. Aggarwal, Philip S. Yu. Method and Apparatus for Analyzing Community Evolution in Graph Data Streams, US Patent 7,890,510, 2011.
16. Charu C. Aggarwal, Philip S. Yu. Method and Apparatus for Query Processing of Uncertain Data. US Patent 7,917,517, 2011.
17. Charu C. Aggarwal. Method, Apparatus and Computer Program Product for Preserving Privacy in Data Mining. US Patent 7,904,471, 2011.

18. Charu C. Aggarwal, Philip S. Yu. Method and Apparatus for Privacy-Preserving Data Mining using Statistical Condensation Approach. US Patent 7,885,941, 2011.
19. Charu C. Aggarwal, Philip S. Yu. Method and apparatus for generating decision trees with discriminants and employing same in data classification, US Patent 7,716,154, 2010.
20. Charu C. Aggarwal. System and Method for Data Traffic Generation via Density Estimation using SVD, US Patent 7,684,963, 2010.
21. Charu C. Aggarwal, Philip S. Yu. Method and apparatus for processing data streams, US Patent 7,739,284, 2010.
22. Charu C. Aggarwal, Philip S. Yu. Apparatus for dynamic classification of data in Evolving Data Stream. US Patent 7,487,167, 2009.
23. Charu C. Aggarwal, Nagui Halim. Method and Apparatus for Privacy-Preserving Data Mining by Restricting Attribute Choice. US Patent 7,475,085, 2009.
24. Charu C. Aggarwal, Philip S. Yu. Method for Dynamic Classification in Evolving Data Stream, US Patent 7,378,939, 2008.
25. Charu C. Aggarwal, Philip S. Yu. Method and Apparatus for Clustering Evolving data Streams through Online and Offline Components, US Patent 7,353,218, 2008.
26. Charu C. Aggarwal. Method and apparatus for flexible data reduction of High Dimensional data, US Patent 7,277,893, 2007.
27. Charu C. Aggarwal. Method and Apparatus for Reduction of High Dimensional Data, US Patent 7,236,638, 2007.
28. Charu C. Aggarwal. Method and Apparatus of Generating Test data sets in accordance with user feedback. US Patent 7,085,981, August 2006.
29. Charu C. Aggarwal. Methods and Apparatus for User-Centered Similarity Learning. US Patent 6,970,884, November 2005.
30. Charu C. Aggarwal, Philip S. Yu. System and Method of Similarity Indexing and Searching in High Dimensional Space. US Patent 6,922,700, July 2005.
31. Charu C. Aggarwal. Method and Apparatus for Classifying Time Series Data using wavelet based approach. US Patent 6,871,165, March 2005.
32. Charu C. Aggarwal. System and Method for Mining Unstructured Data Sets. US Patent 6,847,955, January 2005.
33. Charu C. Aggarwal. Method and Apparatus for User-Centered Class Supervision. 6,804,669, October 2004.
34. Charu C. Aggarwal. System and Method of Determining and Searching for Patterns in A Large Database. US Patent 6,799,175, September 2004.
35. Charu C. Aggarwal, Philip S. Yu. System and Method of Flexible Indexing of Text for Use in Similarity Searches. US Patent 6,785,669, August 2004.
36. Charu C. Aggarwal, Philip S. Yu. System and Method for Classification using Time Sequences. US Patent 6,721,719, April, 2004.
37. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. Method for Targeted Advertising on the Web based on accumulated self learning data, clustering users, and semantic node graph techniques. US Patent 6,714,975, March 2004.
38. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. Method for optimizing profits in electronic delivery of digital objects. US Patent 6,613,413, October 2003.
39. Charu C. Aggarwal, Philip S. Yu. Methods and Apparatus for Performing an Affinity Based Similarity Search. US Patent 6,587,848, July 2003.

40. Charu C. Aggarwal, Philip S. Yu. Method and Apparatus for Similarity Text Search Based on Conceptual Indexing. US Patent 6,542,889, April 2003.
41. Charu C. Aggarwal, Philip S. Yu. Method and Apparatus for Performing Dimensionality Reduction in a Supervised Application Domain. US Patent 6,505,207, January 2003.
42. Charu C. Aggarwal, Joel Wolf, Philip S. Yu. System and Method of Collaborative Filtering with Applications to Ecommerce. US Patent 6,487,541, November 2002.
43. Charu C. Aggarwal, Philip S. Yu. Semantic based collaborative Filtering. US Patent 6,487,539, November 2002.
44. Charu C. Aggarwal, J. Sethuraman, M. Squillante, J. Wolf, P. Yu. Optimizing method for digital content delivery in a multicast network. US Patent 6,477,180, November 2002.
45. Charu C. Aggarwal, Philip S. Yu. System and Method of Using Clustering to Find Personalized Associations. US Patent 6,408,295, June 2002.
46. Ramesh C. Agarwal, Charu C. Aggarwal, V. V. V. Prasad. Depth First Method for Generating Itemsets. US Patent 6,389,416, May 2002.
47. Charu C. Aggarwal, Stephen C. Gates, Philip S. Yu. System and Method for Generating Taxonomies with Applications to Content Based Recommendations. US Patent 6,360,227, March 2002.
48. Charu C. Aggarwal, Philip S. Yu. Content Based Method for Product Peer Filtering. US Patent 6,356,879, March 2002.
49. Charu C. Aggarwal, Philip S. Yu. System and Method for Detecting Clusters of Information with Applications to Ecommerce. US Patent 6,349,309, February 2002.
50. Ramesh C. Agarwal, Charu C. Aggarwal, V. V. V. Prasad. System and Method of Generating Associations. US Patent 6,311,179, October 2001.
51. Charu C. Aggarwal, Joel L. Wolf, Philip Yu. System and Method for Detecting Clusters of Information. US Patent 6,307,965, October 2001.
52. Charu C. Aggarwal, B. Hailpern, Joel L. Wolf, Philip Yu. System and Method for Similarity Searching in High Dimensional Space. US Patent 6,289,354, September 2001.
53. Charu C. Aggarwal, Philip S. Yu. Finding Collective Baskets and Inference Rules for Internet Mining. US Patent 6,263,327, July 2001.
54. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. System and Method for Searching Databases with Applications such as Peer Groups, Collaborative Filtering and Ecommerce. US Patent 6,236,985, May 2001.
55. Charu C. Aggarwal, Philip S. Yu. Methods for Performing Large Scale Auctions and Online Negotiations. US Patent 6,151,589, November 2000.
56. Charu C. Aggarwal, Philip S. Yu. Online Mining of quantitative association rules. US Patent 6,092,064, July 2000.
57. Charu C. Aggarwal, Philip S. Yu. Finding Collective Baskets and inference rules for internet or intranet mining for large databases. US Patent 6,094,645, July 2000.
58. Charu C. Aggarwal, M. Epelman, Joel L. Wolf, Philip S. Yu. System and Method for caching objects of non-uniform size using multiple LRU stack partitions into a range of sizes. US Patent 6,012,126, January 2000.
59. Charu C. Aggarwal, Philip S. Yu. Eliminating Redundancy in Generation of Association Rules for Online Mining. US Patent 5,943,667, August 1999.
60. Charu C. Aggarwal, P. Malkin, B. Schloss, Philip S. Yu. Collaborative caching of a requested object by a lower level node as a function of the caching status of the object a higher level node. US Patent 5,924,116, July 1999.

61. Charu C. Aggarwal, Philip S. Yu. Online Mining of Association Rules. US Patent 5,920,855, July 1999.
62. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. Permutation Based Pyramid Block Transmission Scheme for Broadcasting in Video-On-Demand Storage Systems. US Patent 5,751,336, May 1998.
63. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. System and Method for Construction of a Data Structure for Indexing Multi-Dimensional Objects. US Patent 5,781,906, July 1998.
64. Charu C. Aggarwal, Joel L. Wolf, Philip S. Yu. Maximum Factor Selection Policy for Batching VOD Requests. US Patent 5,631,694, May 1997.

Students Mentored

1. Ramakrishnan Kannan, Georgia Tech., Ph.D. 2016 (Ph.D. thesis committee member)
2. Shiyu Chang, University of Illinois at Urbana-Champaign, Ph.D. 2016 (also Ph.D. thesis committee and supervisor at IBM internship)
3. Jiliang Tang, Arizona State University, Ph.D. 2015 (also Ph.D. thesis committee and supervisor at IBM internship).
4. Guojun Qi, University of Illinois at Urbana-Champaign, Ph.D. 2012 (also Ph.D. thesis committee member and supervisor at IBM internship)
5. Peixiang Zhao, University of Illinois at Urbana-Champaign, Ph.D, 2012 (also Ph.D. thesis committee member and supervisor at IBM internship)
6. Yizhou Sun, University of Illinois at Urbana-Champaign, Ph.D. 2012, (also Ph.D. thesis committee member and supervisor at IBM internship)
7. Min-hsuan Tsai, University of Illinois at Urbana-Champaign, Ph.D. 2013 (also Ph.D. thesis committee member and supervisor at IBM internship)
8. Manish Gupta, University of Illinois at Urbana-Champaign, Ph.D. 2013 (also Ph.D. thesis committee member and supervisor at IBM internship)
9. Dong Wang, University of Illinois at Urbana-Champaign, Ph.D. 2013 (also Ph.D. thesis committee member).
10. Yan Xie, University of Illinois at Chicago, Ph.D. 2012 (also supervisor at IBM internship)
11. Yuchen Zhao, University of Illinois at Chicago, Ph.D. 2012.
12. Michele Dallecheisa, University of Trento at Italy (supervisor at IBM, 2012)
13. Karthik Subbian, University of Minnesota, Ph.D. 2014.

Research Prototypes and Commercialization Experience

- **System S Streaming Fraud Detection Prototype:** Designed and developed a research prototype for fraud detection in fast financial and insurance data streams. Designed the prototype as part of the development process of the IBM Streams product platform (then known as System S during its research design period, and currently a commercial product) for fast processing of incoming data. The approach can be flexibly used for any data mining or streaming application which requires summarization. Demonstrated the capability of the broad algorithm to handle a completely different application such as speaker-recognition. Received **IBM Outstanding Technical Achievement Award** and **IBM Research Division Award** for my work.
- **Streaming Anomaly Detection Prototype:** Designed a number of real time algorithms for abnormality detection in data streams. Designed and developed a prototype implementation which can detect possible alarms for terrorist attacks from medical data streams. The broad technique is also usable for other business intelligence applications. The contributions to this project was recognized by an **IBM Corporate Award for Environmental Excellence** in 2003. Work was patented and also published in SIAM Conference on Data Mining, 2005.

- **Web Portal Personalization Assistant:** My research on text categorization was used to design and develop a system to personalize web portals. As a test case, the particular portal which was implemented using this scheme was the GMAC financial web site of the General Motors Corporation. I also designed and implemented the text categorizer and recommendation system in this IGS offered service. The system provides effective personalization of a web portal, and improves the browsing experience of the customers. Eventually, this approach was generalized and transferred to the *IBM Websphere Portal Product*. The categorization and personalization portions of the work were patented and also published in *IEEE Transactions on Knowledge and Data Engineering*, 16(2), 2004, and *VLDB Conference*, 2002.
- **Interactive Miner:** The prototype was adapted from a research paper in the *ICDE Conference*, 1998. This is a technique for online and interactive generation of association rules with user specified constraints. Uses an OLAP-like preprocess once-query many methodology. Implemented and conceived main ideas behind the prototype. The prototype was released by IBM as an alphaworks release in order to allow the technology to receive public exposure.
- **Profile Miner:** My research on profile association rule mining was used to create this prototype. Designed and developed a prototype, which computes profiles of customers with different kinds of buying behaviors. This approach is useful in e-commerce applications with fast click streams. An example would be an e-commerce merchant with large volumes of transactional behavior. This approach provides interactive mining ability with the use of a preprocess-one query-many paradigm. The prototype was released by IBM as part of its alphaworks suite of releases. Work was patented and published in *KDD Conference*, 1998.
- **Web Traffic Express 1.1 Caching Algorithm:** Designed a caching algorithm for web objects. This was one of the earliest algorithms for web object caching. An algorithm was designed, which takes the sizes of the underlying data objects into account. The resulting algorithm was transferred to *Web Traffic Express 1.1*. and was published in *IEEE TKDE 1999*.

Journal Editorships and Editorial Boards

- Associate editor, *IEEE Transactions on Knowledge and Data Engineering*, 2003-2007.
- Associate Editor, *ACM Transactions on Knowledge Discovery from Data*, 2011–present
- Action Editor, *Data Mining and Knowledge Discovery Journal*, 2005-present
- Advisory Board Member, Springer, *Lecture Notes in Social Networks*, 2013.
- Editorial Board, *Journal of Database Management Research*, 2009-2010, 2012-present.
- Editorial Board, *Knowledge and Information Systems*, 2009-present.
- Editor-in-Chief, *ACM SIGKDD Explorations*, 2014-present.

Conference Organization, Keynotes, and Other Professional Service

- Program co-chair of ACM KDD Conference, 2016.
- Program co-chair at IEEE ICDM Conference, 2015.
- Program co-chair of Knowledge Management Track, *ACM Conference on Information and Knowledge Management (CIKM)*, 2015.
- General co-chair of IEEE Big Data Conference, 2014.
- Program co-chair of Data Mining Track, *Worldwide Web (WWW) Conference*, 2009.
- Keynote Speaker at the ECML/PKDD Conference, 2014.
- Keynote Speaker at the ASONAM Conference, 2014.
- Keynote Speaker at the European Conference on Machine Learning, 2006.

- Vice-President, SIAM Activity Group on Data Mining (this activity group is responsible for running all data mining activities at SIAM), 2011-2013.
- Member of **SIAM Industry Committee**
- Member of **IEEE ICDM Steering Committee**
- Regularly serve as senior program committee member or program committee member at major data mining and database conferences.

Citizenship

- United States Citizen