Kanthi K Sarpatwar

🏠: Briarcliff Manor, New York, ✉: kanthik@gmail.com, ☎: (408) 707-6708

Professional Interests

Machine Learning and Optimization: Differential Privacy, Homomorphic Encryption, Bayesian Optimization, Explainability for AI and Optimization
Algorithms & Theory: Design and Analysis of Algorithms for NP-Hard Optimization problems; Combinatorial Optimization Techniques; Submodular Optimization; Scheduling Problems

Education

University of Maryland, College Park, MD, USA — **Advisor:** Samir Khuller
Ph.D. in Computer Science — Aug 2010 - May 2015
Doctoral Dissertation Title: Allocation Algorithms for Networks with Scarce Resources
Indian Institute of Technology, Madras, India — **Advisor**: Narayanaswamy N. S.
M.Tech. in Computer Science — Jul 2003 - May 2008
B.Tech. in Computer Science — Jul 2003 - May 2008

Professional Appointments

Senior Research Scientist — *Jul 2022-Present*
IBM Research AI, T.J. Watson Center Yorktown Heights, NY
Research Staff Member — *May 2015-Jul 2022*
IBM Research AI, T.J. Watson Center Yorktown Heights, NY
Software Analyst — *Jun 2008- Apr 2010*
Global Analytics, Chennai, India
Summer Research Internship — *Jun-Aug 2014*
IBM Research AI, T.J. Watson Center, Yorktown Heights, NY
Summer Research Internship — *May-Jun 2014*
Rutgers University, Camden, NJ
Summer Research Internship — *Jun-Aug 2013*
Alcatel-Lucent Bell Labs (Currently Nokia Bell Labs), Murray Hill, NJ
Summer Software Internship — *Jun-Aug 2012*
Yahoo Inc., Santa Clara, CA

Selected Awards & Honors

Research Accomplishment Award, IBM Research. — 2021.
For contributions in privacy preserving machine learning applied to IBM Z-systems.
Multiple (2) Equity Awards, IBM. — 2018 and 2021.
This is an annual award given to top 1% performing IBM employees globally.
Multiple (4) Invention Plateau Awards, IBM — 2018, 2019, 2020, 2021.
IBM Manager's Choice Award, IBM Research — 2018.
Future Faculty Fellow. University of Maryland, College Park. — 2014-2015.
Dean's Graduate Fellowship Award. University of Maryland, College Park. — 2010-12.
MCM Scholarship. Government of India. — 2003-2007.
Prathiba Scholarship. State of Andhra Pradesh, India. — 2003.

Selected Recent Research Projects

Tech Lead: Privacy Preserving Machine Learning — 2019-2021
Contributed to the development of a library for *secure machine learning inferencing on encrypted data*. This library played a critical role in a demonstration on privacy preserving credit card fraud detection at **IBM Think Conference in 2021** and won an **IBM research accomplishment** award. Leveraging advanced privacy-techniques, specifically in homomorphic encryption, and differential privacy, and combining them with novel machine learning ideas the library supported several machine learning algorithms: sparse linear regression techniques (such as lasso, ridge), logistic regression, support vector machines, shallow neural networks, decision tree regressor/classifier, ensembles of decision trees (such

as random forest, adaboost, gradient boosting), k-means clustering, mixture models and isolation forests.

Tech Lead: Explainability for AI 2021-

Contributed to the development of various explainability capabilities based on global/local feature importance, conditional analyses, SHAP-based explanations and contrastive explanations for blackbox machine learning models dealing with classification, regression and anomaly detection tasks. These capabilities are now part of various IBM product offerings.

Tech Lead: Explanations for Large Scale Blackbox Schedule Optimizers 2021 -

Contributed to the development of a state of the art, first of its kind explanation capability aimed to help industry professionals to better understand complex scheduling decisions. Developed and implemented a novel combinatorial algorithm with provable quality guarantees in terms of explanation complexity (2-approximation guarantee compared to the optimal complexity) and with a runtime improvement of 20x over the state of the art approaches. This capability has become part of IBM product offerings.

Bayesian Optimization for Inventory Optimization 2020-2021

Key contributor in the development of a novel Bayesian optimization approach for optimizing a multi-echelon inventory system. Dealing with dependencies across different locations of the system, in a traditional way, requires one to compute complex convolutions of various demand/supply distributions. We by-passed these complexities by employing a blackbox approach which exploited state-of-the-art Bayesian optimization techniques and libraries (such as BO-Torch) and lead to roughly 40% improvement over pre-existing product approach.

Publications

Testing Membership in Distributional Simplex
G. Ganapavarapu, **K. Sarpatwar**, K. Shanmugam.
Manuscript under prep/submission

Optimizer Agnostic Explainability for Large Scale Schedules
S. S. K. Sajja, **K. Sarpatwar**, L. M. Nguyen, Y.Y. Jia, S. Michel, R. Vaculin.
Manuscript under prep/submission

High Throughput Machine Learning Inference on Homomorphically Encrypted Data
K. Sarpatwar, E. Ahorani, J. Rayfield, H. Min, H. Shaul, E. Kushnir, O. Soceanu, D. Dillenberger, R. Vaculin.
Manuscript under prep/submission

FHE-Friendly Distillation of Decision Tree Ensembles for Efficient Encrypted Inference
K. Nandakumar, **K. Sarpatwar**, N. Ratha, S. Pankanti, R. Vaculin, K. Shanmugam, J. Rayfield
Manuscript under prep/submission. Preliminary version accepted as an **Oral** *talk at ACM CCS 2021 Workshop on Privacy Preserving Machine Learning*

Peer-reviewed International Conferences

24. Maximizing Throughput in Flow Shop Real-time Scheduling
L. B. Yamin, J. Li, **K. Sarpatwar**, B. Schieber and H. Shachnai.
International Conference on Approximation Algorithms for Combinatorial Optimization Problems (APPROX 2020).

23. Privacy Enhanced Decision Tree Inference K. Sarpatwar, N. Ratha, K. Nandakumar, K. Shanmugam, J. Rayfield, S. Pankanti, R. Vaculin

IEEE CVPR Workshop on Fair, Data Efficient and Trusted Computer Vision, 2020

21. The Preemptive Resource Allocation Problem.
K. Sarpatwar, B. Schieber and H. Shachnai.
39th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2019).

21. Differentially Private Distributed Data Summarization under Covariate Shift
K. Sarpatwar, K. Shanmugam, A. Jagmohan, R. Vaculin and G. Ganapavarapu.
Thirty-third Conference on Neural Information Processing Systems (NeurIPS 2019)

20. Blockchain Enabled AI Marketplace: The Price You Pay For Trust
K. Sarpatwar, G. Ganapavarapu, K. Shanmugam, A. Rahman, R. Vaculin.
Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019.

19. Generalized Assignment with Group Constraints.
A. Kulik, **K. Sarpatwar**, B. Schieber and H. Shachnai
European Symposium on Algorithms (ESA 2019).

18. Online Resource Allocation with Matching Constraints.
J. P. Dickerson, K. A. Sankararaman, **K. Sarpatwar**, A. Srinivasan, K. L. Wu, P. Xu.
International Conference on Autonomous Agents and Multiagent Systems (AAMAS), 2019.

17. Brief Announcement: Approximation Algorithms for Preemptive Resource Allocation.
K. Sarpatwar, B. Schieber and H. Shachnai
Symposium on Parallelism in Algorithms and Architectures (SPAA), 2018.

16. Generalized Assignment of Time-Sensitive Item Groups.
K. Sarpatwar, B. Schieber and H. Shachnai
In the 21st International Conference on Approximation Algorithms for Combinatorial Optimization Problems (APPROX), 2018.

15. Budgeted Online Assignment in Crowdsourcing Markets: Theory and Practice.
P. Xu, A. Srinivasan, **K. Sarpatwar** and K.L. Wu
In the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS), 2017.

14. Approximation Algorithms for Connected Maximum Cut and Related Problems.
M.T. Hajiaghayi, G. Kortsarz, R. MacDavid, M. Purohit, and **K. Sarpatwar**.
In the 23rd Annual European Symposium on Algorithms (ESA), 2015.

13. The Container Selection Problem.
V. Nagarajan, **K. Sarpatwar**, B. Schieber, H. Shachnai, and J. L. Wolf
In the 18th International Conference on Approximation Algorithms for Combinatorial Optimization Problems (APPROX), 2015.

12. The X-flex Cross Platform Scheduler: Who's the Fairest of Them All?
J. Wolf, Z. Nabi, V. Nagarajan, R. Saccone, R. Wagle, K. Hildrum, E. Pring, and **K. Sarpatwar**.
In the 15th ACM/IFIP/USENIX Middleware 2014 - Industry Track.

11. Analyzing the Optimal Neighborhood: Algorithms for Budgeted and Partial Connected Dominating Set Problems.
S. Khuller, M. Purohit, and **K. Sarpatwar**
In the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2014.

10. New Approximation Results for Resource Replication Problems.
S. Khuller, B. Saha, and **K. Sarpatwar**
In the 15th International Conference on Approximation Algorithms for Combinatorial Optimization Problems (APPROX) 2012.

9. Rainbow Connectivity: Hardness and Tractability.
P. Ananth, M. Mande, and **K. Sarpatwar**.
IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS), 2011.

Journals

8. The Preemptive Resource Allocation Problem.
K. Sarpatwar, B. Schieber and H. Shachnai.
Journal of Scheduling, 2021 (accepted after minor revision).

7. Analyzing the Optimal Neighborhood: Algorithms for Partial and Budgeted Connected Dominating Set Problems.
S. Khuller; M. Purohit; **K. Sarpatwar**
SIAM Journal of Discrete Mathematics, 2020.

6. Approximation Algorithms for Connected Maximum Cut and Related Problems.
M.T. Hajiaghayi; G. Kortsarz; R. MacDavid; M. Purohit; **K. Sarpatwar**
Theoretical Computer Science, 2019 (Accepted with Minor Revision).

5. Constrained Submodular Maximization via Greedy Local Search.
K. K. Sarpatwar, B. Schieber and H. Shachnai.
Operations Research Letters, 2019

4. On Maximum Leaf Trees and Connections to Connected Maximum Cut Problems.
R. Gandhi, M. T. Hajiaghayi, G. Kortsarz, M. Purohit, **K. K. Sarpatwar**
Information Processing Letters (2017).

3. New Approximation Results for Resource Replication Problems.
S. Khuller, B. Saha, and **K. K. Sarpatwar**
Algorithmica (2015)

2. Blockchain Analytics and Artificial Intelligence.
D.N. Dillenberger, P. Novotny, Q. Zhang, P. Jayachandran, H. Gupta, S. Hans, D. Verma, S. Chakraborty, J.J. Thomas, M.M. Walli, R. Vaculin, **K. Sarpatwar**
IBM Journal of Research and Development, 2019

1. Hardness of Subgraph and Supergraph Problems in r-tournaments.
K. K. Sarpatwar, N. S. Narayanaswamy
Theoretical Computer Science (2011)

Book Chapters

Towards Enabling Trusted Artificial Intelligence via Blockchain.

K. K. Sarpatwar, R. Vaculin, H. Min, G. Su, T. Heath, G. Ganapavarapu, D. Dillenberger. Policy-Based Autonomic Data Governance. Springer 2019.

Invention Disclosures & Patents

22. **System and Method for Handling Missing Pertinent Negatives in Generating Contrastive Explanations.** Sumanta Mukherjee, Kanthi Sarpatwar, Sattwati Kundu, Raghunath E Nair, Shraddha Singh, Roman Vaculin. Accepted to File

21. **Improving the Execution of Neural Network Inference under Homomorphic Encryption through Efficient Operation Elimination and Grouping.** Subhankar Pal, Alper Buyuktosunoglu, Ehud Aharoni, Nir Drucker, Omri Soceanu, Hayim Shaul, Kanthi Sarpatwar, Roman Vaculin. Filed

20. **Configurable neighborhood generation for local explainability.** Natalia Martinez Gil, Kanthi Sarpatwar, Roman Vaculin. Accepted to be filed

19. **Optimizer Agnostic Explainability for Large Scale Schedules.** Surya Shravan Kumar Sajja, Kanthi Sarpatwar, Lam M. Nguyen, Yuan Yuan Jia, Stephane Michel, Roman Vaculin. *Patent Filed.*

18. **Testing Membership in Distributional Simplex**. Giridhar Ganapavarapu, Kanthi Sarpatwar, Karthikeyan Shanmugam. *Accepted to be filed.*

17. **Efficiently Batching Pre-Encrypted Data for Homomorphic Inference.** Kanthi Sarpatwar, Roman Vaculin, Ehud Aharoni, James T. Rayfield, Omri Soceanu. *Patent Filed.*

16. **Multi-Phase Privacy-Preserving Inferencing in a High Volume Data Environment**. Roman Vaculin, Kanthi Sarpatwar, Hong Min. *Patent Filed*

15. **FHE Friendly Knowledge distillation**. Kanthi Sarpatwar, Nalini Ratha, Karthikeyan Shanmugam, Karthik Nandakumar, Sharathchandra Pankanti, James T Rayfield and Roman Vaculin. *Patent Filed.*

14. **System and Method for Privacy Enhanced Decision Tree based Classification.** Kanthi Sarpatwar, Nalini Ratha, Karthikeyan Shanmugam, Karthik Nandakumar, Sharathchandra Pankanti and Roman Vaculin. *Patent Filed.*

13. **Decision Tree-Based Inference on Homomorphically Encrypted Data without Bootstrapping**. Kanthi Sarpatwar, Nalini Ratha, Karthikeyan Shanmugam, Karthik Nandakumar, Sharathchandra Pankanti and Roman Vaculin. *Patent Filed*

12. **System and Method for Private Verification of Functional Intellectual Property**. Giridhar Ganapavarapu, Kanthi Sarpatwar, Roman Vaculin, Gi-Joon Nam. *Defensive Publication.*

11. **A General System and Method for Enabling Boosting Protocols on Encrypted Data**. Kanthi Sarpatwar, Roman Vaculin. *Patent Filed.*

10. **A System and Method for Efficient Unsupervised Anomaly Detection on Encrypted Data**. Kanthi Sarpatwar, Venkata Sitaramagiridharganesh Ganapavarapu, Saket Sathe, Roman Vaculin. *Patent Granted #US11271958B2*

9. **System and Method for Private Verification of Stochastic Gradient Descent**. Giridhar Ganapavarapu, Kanthi Sarpatwar, Karthikeyan Shanmugam, Roman Vaculin. *Patent Filed.*

8. **Efficient Database Machine Learning Verification**. Giridhar Ganapavarapu, Kanthi Sarpatwar, Karthikeyan Shanmugam, Roman Vaculin. *Patent Filed*

7. **Aggregated Machine Learning Verification for Databases**. Giridhar Ganapavarapu, Kanthi Sarpatwar, Karthikeyan Shanmugam, Roman Vaculin. *Patent Filed.*

6. **Efficient Verification of Machine Learning Applications**. Giridhar Ganapavarapu, Kanthi Sarpatwar, Karthikeyan Shanmugam, Roman Vaculin. *Patent Filed*

5. **Artificial Intelligence Software Marketplace**. K. K. Sarpatwar, K. Shanmugam, A. Jagmohan, M. M. Franceschini, R. Vaculin. *Patent App: US20190287027A1*

4. **Multi-platform Scheduler for Permanent and Transient Applications**. P. Kirchner, K. Onak, R. Saccone, K. K. Sarpatwar, J. Wolf. Patent Granted: US20180081722A1.

3. **Cross Platform Scheduling with Long-term Fairness and Platform Specific Optimization.** K. Hildrum, Z. Nabi, V. Nagarajan, R. Saccone, K. K. Sarpatwar, R. Wagle, J. Wolf. Patent Granted# US9886307B2.

2. Cross Platform Scheduling with Long-term Fairness and Platform Specific Optimization. K. Hildrum, Z. Nabi, V. Nagarajan, R. Saccone, K. K. Sarpatwar, R. Wagle, J. Wolf. Patent Granted# US9886307B2.
1. System and Method for Fast Network Queries. R. Bhatia, B. Gupta, K. K. Sarpatwar, L. Greenwald. US Patent# US9886306B2.

Skills

Analytical & Mathematical Skills: Expert in the design and analysis of algorithms (particularly related to ML problems), experienced in mathematical programming/modeling
Programming Languages: 10 years experience with Python, C++, and C. 5 years of enterprise ML engineering experience.
Machine Learning Tools: Scikit-learn, TensorFlow, PyTorch, BoTorch, Numpy, Scipy, Pandas.
Database Scripting. MySQL, CouchDB.
Operating Systems. Linux, DOS and OSX.
Privacy-preserving Tools. Homomorphic Encryption libraries IBM HElib, HEAAN; Differential privacy techniques
Mathematical Programming Tools. CPLEX.

Talks Presented

Explainability For Efficient And Trusted Decision Optimization
9. INFORMS Annual Meeting 2021.
Trusted AI Marketplaces with Privacy Guarantees: A Tale of Two Hashes
8. IBM Blockchain Research Directions Worldwide Call.
Preemptive Resource Constrained Scheduling with Time-Windows.
7. At DIMACS Workshop on Algorithms for Data Center Networks (Jun 5-7, 2017).
Generalized Assignment of Time-Sensitive Item Groups
6. APPROX 2018, Princeton University.
The Container Selection Problem
5. APPROX 2015, Princeton University.
4. Intern Exit Talk, IBM TJ Watson Research, New York, Aug 2014.
Analyzing the Optimal Neighborhood: Algorithms for Budgeted and Partial Connected Dominating Set Problems.
3. SODA 2014. Portland, Oregon, Jan 7, 2014.
Approximate Oracle for Answering Fundamental Graph Queries.
2. Intern Exit Talk, Bell Labs, Aug 2013.
Indexing as a Service
1. Intern Exit Talk, Yahoo Inc., Aug 2012.

Mentoring

Pan Xu, University of Maryland (College Park). *Current Position: Assistant Professor in the Department of Computer Science at New Jersey Institute of Technology*. **Topic:** Online Resource Allocation with Matching Constraints.
Akond Rahman, North Carolina State University. *Current Position: Assistant professor in the Department of Computer Science (CSC) at Tennessee Tech University*. **Topic:** Blockchain Enabled AI Marketplace
Mayank Saxena, Columbia University. **Topic:** Scalable Unsupervised Learning Algorithms on Homomorphically Encrypted Data.
Natalia Lucienne Martinez Gil, Duke University. **Topic:** Configurable neighborhood generation for local explainability

Review Services

Program Committee Member AAAI 2022, AISTATS 2022, ICML 2022, UAI 2022
Program Committee Member AISTATS 2021, NeurIPS 2021, ICML 2021, UAI 2021.
Program Committee Member IJCAI 2020, ICML 2020, NeurIPS 2020.
Reviewer for AIStats 2020, STACS 2020, SWAT 2020.
Program Committee Member, IJCAI 2019.

Integer Programming and Combinatorial Optimization – 2019
IEEE/ACM Transactions on Networking – 2019
Symposium on Discrete Algorithms (SODA) – 2014, 2015, 2016, 2017, 2018
International Colloquium on Automata, Languages and Programming (ICALP) – 2018
Symposium on Parallel Algorithms and Architectures (SPAA) – 2015
IEEE International Conference on Computer Communications (INFOCOM) – 2015
Journal of Discrete Optimization – 2014
Journal of Algorithms – 2017
International Computing and Combinatorics Conference (COCOON) – 2016
Journal of Mathematics of Operations Research – 2017
Information Processing Letters – 2015
IEEE Transactions on Parallel and Distributed Systems – 2015
Symposium on Theoretical Aspects of Computer Science – 2015
European Symposium on Algorithms – 2015
Foundations of Software Technology and Theoretical Computer Science (FSTTCS) – 2014

References Available on request.