

# Saurav Prakash

Ming Hsieh Department of Electrical and Computer Engineering,  
University of Southern California, Los Angeles, CA 90089

✉ sauravpr@usc.edu  
🌐 <https://sauravpr.com/>

## Education

2016 – **Ph.D. in Electrical Engineering.**

Present *University of Southern California*, Los Angeles, CA, USA.

Advisor: Prof. Salman Avestimehr.

2012 – 2016 **B.Tech. in Electrical Engineering.**

with Minor in Artificial Intelligence in Computer Science and Engineering.

*Indian Institute of Technology Kanpur*, Kanpur, UP, India.

Advisor: Prof. Aditya K. Jagannatham.

## Research Interests

- Security and Privacy in Machine Learning
- Efficient and Robust Federated Learning
- Large-Scale Decentralized Serverless Training
- Coded Distributed Computing
- Information and Coding Theory

## Research Overview

I have worked extensively towards holistically addressing real-world bottlenecks in large-scale distributed computing, including both *large-scale cloud computing* and *decentralized machine learning*. My contributions involve novel solutions that target multiple challenges in both these broad domains, bringing new concepts from Shannon's information theory, coding theory, communication topology design, data privacy, and optimization theory.

## Professional Experience

Aug. 2016 – **Graduate Research Assistant.**

Present *vITAL Lab, University of Southern California*, Los Angeles, CA.

Mentor: Prof. Salman Avestimehr.

- Byzantine Robust Federated Learning
  - Considered general Byzantine federated learning setting with non-IID data across clients
  - Proposed DiverseFL, a novel sampling based approach that applies per client criteria for mitigating Byzantines in the general federated learning setting
  - Demonstrated via extensive experiments that compared to prior approaches, DiverseFL performs much better, almost achieving the optimal model performance

- Byzantine Robust Decentralized (Serverless) Learning
  - Considered the problem of Byzantine mitigation in the decentralized learning setting without any central coordinator
  - Developed Basil, a fast and computationally efficient Byzantine robust algorithm leveraging a sequential, memory assisted and performance criteria for training over a ring
  - Demonstrated numerically that Basil provides up to  $\sim 16\%$  higher test accuracy when compared with prior methods, under different Byzantine fault settings
- Coded Computing for Hierarchical Distributed Learning at the Edge
  - Formulated a hierarchical gradient aggregation problem for machine learning from data available at the client nodes by leveraging reliable helper nodes for collecting updates
  - Proposed two unique coded computing strategies – aligned repetition coding (ARC), aligned MDS coding (AMC) – for mitigating straggling links from clients to helpers
- Coded Computing for Large-scale Distributed Learning
  - Formulated a tree gradient coding framework and proposed CodedReduce scheme for fast and robust gradient aggregation in distributed learning
  - CodedReduce combines advantages of communication efficiency of Ring-AllReduce and straggler resiliency of Gradient Coding for minimizing the overall training latency
  - In experiments over Amazon EC2, CodedReduce achieves gains of up to  $31\times$  in the overall execution time over prior approaches for distributed learning
- Coded Computing for Large-scale Graph Processing
  - Proposed a distributed computing framework for graph analytics based on MapReduce
  - Characterized the optimal trade-off between Map computations and Shuffle load for the Erdos-Renyi model
  - Developed and implemented a coded distributed implementation of the PageRank algorithm using Amazon EC2, demonstrating gains of up to 50% over the naive PageRank
  - Developed coding schemes for three other popular random graph models – random bi-partite model, stochastic block model, and power law model
- Coded Computing for Large-scale Matrix Multiplication in Heterogeneous Settings
  - Proposed a two-step alternative formulation to the problem of minimizing the expected run-time in distributed matrix-vector multiplication in heterogeneous clusters
  - Developed a scalable method – Heterogeneous Coded Matrix Multiplication (HCMM) – for reliable matrix multiplication on cloud clusters with stragglers
  - Proved the asymptotic optimality of HCMM
  - Implemented HCMM using Amazon EC2, demonstrating gains of up to 61% over benchmark schemes

May – Aug. **Graduate Technical Intern.**

2018 & 2019 *Intel Labs*, Santa Clara, CA.

Mentors: Sagar Dhakal, Nageen Himayat, Shilpa Talwar.

- Coded Computing for Federated Learning in Multi-access Edge Computing (MEC) networks
  - Proposed CodedFedL for injecting coding redundancy into federated learning with non-IID data for mitigating stragglers and minimizing up training time in MEC networks
  - Developed a tractable approach for minimizing deadline time
  - Analyzed the convergence rate and iteration complexity of CodedFedL
  - Demonstrated gains of up to  $15\times$  in comparison to benchmark schemes in practice

Jun. 2021 – **Applied Scientist Intern.**

Aug. 2021 *Alexa AI, Amazon*, Cambridge, MA.

Mentors: Clement Chung, Christophe Dupuy, Rahul Gupta, Leo Long, Tanya Roosta.

- Federated learning with Heterogeneous Model Architectures
  - Developed various strategies for efficient federated learning from edge users
  - Explored novel methods to enable federated learning with heterogeneous model architectures at the edge users

May 2015 – **International Visiting Student.**

Jul. 2015 *IUSSTF-Viterbi Program*, Los Angeles, CA.

Mentor: Prof. Salman Avestimehr.

- Towards Faster Algorithms for Processing Large Data on Graphs
  - Studied spectral graph theory and its application in signal processing of graph data – cut-off frequency, optimal sampling and bandlimited interpolation
  - Explored existing semi-supervised and active learning methods for data on graphs
  - Proposed a Random Jump model based on Graph Laplacian for sampling with low time complexity

2013 – 2016 **Undergraduate Research Assistant.**

*MWN Group, IIT Kanpur*, Kanpur, India.

Mentor: Prof. Aditya K. Jagannatham.

- Scheduling for Efficient Utilization of Time Resource in Wireless Networks
  - Worked on the problem of user scheduling for efficient wireless resource utilization, under resource allocation fairness constraints
  - Proposed two opportunistic schemes for scheduling users in a time slotted system with wireless Rayleigh-fading channel
  - Simulations predicted stochastically improved performance compared to Round Robin scheme alongside satisfaction of any arbitrary time resource allocation fairness constraints

May 2013 – **Undergraduate Research Intern.**

Jul. 2013 *Summer Undergraduate Research Grant for Excellence (SURGE)*, IIT Kanpur, Kanpur, India.

Mentor: Prof. Aditya K. Jagannatham.

- Channel Estimation and Capacity in MIMO Wireless Communication Systems
  - Studied capacity lower bound for a MIMO system obeying Block-Fading law using LMMSE estimator for channel estimation at the receiver
  - Simulated existing algorithms designed for obtaining near optimal number of transmit antennas for optimizing the capacity lower bound

---

## Selected Honors and Awards

2021 Qualcomm Innovation Fellowship.

2019 Qualcomm Innovation Fellowship Finalist.

2019 Most Novel Research Project Award, EE-599 (Deep Learning course, USC).

2016 USC Annenberg PhD Fellowship.

- 2016 Princeton Gordon Wu PhD Fellowship Offer.
- 2015 Viterbi-India Internship.
- 2014 Summer Undergraduate Research Grant for Excellence at IIT Kanpur.
- 2015 Shri Singhasan Singh Scholarship at IIT Kanpur.
- 2015 Institution of Engineering and Technology (IET) Scholarship.

## Publications

Link to Google Scholar.

(\* denotes joint authorship).

### Preprints

- P2 **S. Prakash**, H. Hashemi, Y. Wang, M. Annavaram, S. Avestimehr, “Byzantine resilient federated learning with heterogeneous data distribution,” Jul. 2021. *Partly presented at the Enclaved AI/ML Workshop 2021, Private AI Research Institute.*
- P1 A. R. Elkordy, **S. Prakash**, S. Avestimehr, “Basil: A fast and Byzantine-resilient approach for decentralized training,” Sep. 2021. *Part of it to be presented at the NeurIPS Workshop on Privacy in Machine Learning, 2021.*

### Journal Papers

- J4 **S. Prakash**, S. Dhakal, M. Akdeniz, Y. Yona, S. Talwar, S. Avestimehr, N. Himayat, “Coded computing for low-latency federated learning over wireless edge networks,” *IEEE Journal on Selected Areas in Communications*, volume 39, issue 1, pages 233–250, Jan. 2021. *Was partly presented at the FL-ICML Workshop on User Privacy and Data Confidentiality, 2020.*
- J3 A. Reisizadeh\*, **S. Prakash\***, R. Pedarsani, S. Avestimehr, “Coded Reduce: A fast and robust framework for gradient aggregation in distributed learning,” to appear in the *IEEE/ACM Transactions on Networking*.
- J2 **S. Prakash\***, A. Reisizadeh\*, R. Pedarsani, S. Avestimehr, “Coded computing for distributed graph analytics,” *IEEE Transactions on Information Theory*, volume 66, issue 10, pages 6534–6554, Oct. 2020 .
- J1 A. Reisizadeh, **S. Prakash**, R. Pedarsani, S. Avestimehr, “Coded computation over heterogeneous clusters,” *IEEE Transactions on Information Theory*, volume 65, issue 7, pages 4227–4242, Jul. 2019 .

### Conference/Workshop Proceedings

- C7 **S. Prakash\***, A. Reisizadeh\*, R. Pedarsani, S. Avestimehr, “Hierarchical coded gradient aggregation for learning at the edge,” in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Aug. 2020 .
- C6 S. Dhakal, **S. Prakash**, Y. Yona, S. Talwar, N. Himayat, “Coded federated learning,” in *Proceedings of IEEE Globecom Workshops (GC Wkshps)*, Mar. 2020.
- C5 S. Kundu\*, **S. Prakash\***, H. Akrami, P. Beerel, K. Chugg, “pSConv: A pre-defined sparse kernel based convolution for deep CNNs,” in *Proceedings of IEEE 57<sup>th</sup> Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Dec. 2019.

- C4 S. Dhakal\*, **S. Prakash\***, Y. Yona, S. Talwar, N. Himayat, “Coded computing for distributed machine learning in wireless edge network,” in *Proceedings of IEEE 90<sup>th</sup> Vehicular Technology Conference (VTC2019-Fall)*, Nov. 2019.
- C3 A. Reisizadeh\*, **S. Prakash\***, R. Pedarsani, S. Avestimehr, “Tree gradient coding,” in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Sep. 2019.
- C2 **S. Prakash\***, A. Reisizadeh\*, R. Pedarsani, S. Avestimehr, “Coded computing for distributed graph analytics,” in *Proceedings of IEEE International Symposium on Information Theory (ISIT) Conference*, Aug. 2018.
- C1 A. Reisizadeh, **S. Prakash**, R. Pedarsani, S. Avestimehr, “Coded computation over heterogeneous clusters,” in *Proceedings of IEEE International Symposium on Information Theory (ISIT) Conference*, Aug. 2017.

## Selected Talks

- Jul. 2021 **TEE-GPU Cooperative Learning: Privacy and Security Without the Price**, *Presentation, Enclaved AI/ML Workshop 2021*, Private AI Research Institute.
- May 2021 **Federated deep learning: On-device learning for CV and NLP**, *Finalist Team Presentation, Qualcomm Innovation Fellowship 2021*, Qualcomm.
- Apr 2021 **Trustworthy and Scalable Federated Learning**, *CCF Advanced Disciplines Lecture*, Institute of Computing Technology, Chinese Academy of Sciences.
- Jul. 2020 **Coded Computing for Federated Learning at the Edge**, *Presentation, FL Workshop on User Privacy and Data Confidentiality*, International Conference on Machine Learning.

## Patents

- 2021 M. R. Akdeniz, A. Anand, N. Himayat, A. S. Avestimehr, R. Balakrishnan, P. Bhardwaj, J. Choi, Y.-S. Choi, S. Dhakal, B. G. Edwards, **S. Prakash**, A. Solomon, S. Talwar, Y. E. Yona, “Systems and methods for distributed learning for wireless edge dynamics,” *App. No. PCT/US2020/067068*.
- 2019 **S. Prakash**, S. Dhakal, Y. Yona, N. Himayat, S. Talwar, “Technologies for distributing iterative computations in heterogeneous computing environments,” *US Patent App. 16/368,716*.
- 2019 **S. Prakash**, S. Dhakal, Y. Yona, N. Himayat, S. Talwar, “Technologies for distributing gradient descent computation in a heterogeneous multi-access edge computing (MEC) networks,” *US Patent App. 16/235,682*.

## Professional Service

- 2017–2021 **Invited Journal Reviewer.**
  - IEEE Journal on Selected Areas in Communications
  - IEEE Transactions on Information Theory
  - IEEE Journal on Selected Areas in Information Theory
  - IEEE Transactions on Communications
- 2017–2021 **Invited Conference/Workshop Reviewer.**
  - IEEE International Symposium on Information Theory (ISIT)
  - IEEE Information Theory Workshop (ITW)

## ■ Mentorship Experience

- 2021 **Mentor**, *Graduate Application Mentorship Program (GradAMP)*, USC.
- 2013 – 2016 **Student Guide**, *Counselling Service*, IIT Kanpur.
- 2013 – 2015 **Ambassador Caller**, *Alumni Contact Program*, IIT Kanpur.
- 2013 – 2014 **Student Secretary**, *Fine Arts Club*, IIT Kanpur.

## ■ Selected Coursework

### **Algorithms and Artificial Intelligence.**

- Deep Learning
- Data Structure and Algorithms
- Fundamentals of Computing
- Machine Learning for Computer Vision
- Artificial Intelligence Programming

### **Mathematics.**

- Real Analysis
- Probability and Statistics
- Linear Algebra
- Partial Differential Equations
- Random Processes in Engineering

### **Communication and Signal Processing.**

- Error Correcting Codes
- Wireless Communications
- Information Theory
- Topics in Cryptography and Coding
- Digital Signal Processing
- Digital Communication Networks