

# Saurav Prakash

---

## CONTACT INFORMATION

EEB 514, 3740 McClintock Ave,  
Electrical and Computer Engineering,  
University of Southern California,  
Los Angeles, CA 90089.  
Email: [sauravpr@usc.edu](mailto:sauravpr@usc.edu)  
Phone: +1 5714242398

[Google Scholar](#)  
[LinkedIn](#)

## RESEARCH INTERESTS

Secure and Privacy-Preserving Machine Learning, Federated Learning, Large-Scale Serverless Training, Coded Distributed Computing, Information and Coding Theory

## EDUCATION

*Doctor of Philosophy* in Electrical and Computer Engineering  
University of Southern California

*Aug 2016 – present*

*Bachelor of Technology* in Electrical Engineering

*Jul 2012 – May 2016*

*Minor* in Artificial Intelligence in Computer Science and Engineering  
Indian Institute of Technology Kanpur

## PUBLICATIONS

### Preprints and Journal Articles

- Saurav Prakash, Hanieh Hashemi, Yongqin Wang, Murali Annavaram and A. Salman Avestimehr. “[Byzantine-resilient federated learning with heterogeneous data distribution](#),” e-print.
- Ahmed Roushdy Elkordy, Saurav Prakash, A. Salman Avestimehr. “[Basil: A fast and Byzantine-resilient approach for decentralized training](#),” e-print .
- Saurav Prakash, Sagar Dhakal, Mustafa Akdeniz, Yair Yona, Shilpa Talwar, Salman Avestimehr and Nageen Himayat. “[Coded computing for low-latency federated learning over wireless edge networks](#),” IEEE Journal on Selected Areas in Communications, 2021.
- Saurav Prakash, Sagar Dhakal, Mustafa Akdeniz, A. Salman Avestimehr, and Nageen Himayat. “[Coded computing for federated learning at the edge](#),” e-print. Presented at the FL-ICML Workshop on User Privacy and Data Confidentiality, 2020.
- Amirhossein Reisizadeh\*, Saurav Prakash\*, Ramtin Pedarsani, and Salman Avestimehr. “[Coded Reduce: A fast and robust framework for gradient aggregation in distributed learning](#),” IEEE/ACM Transactions on Networking, 2021. Partly presented at the ICML Workshop on Coding Theory for Large-Scale Machine Learning, 2019.
- Saurav Prakash\*, Amirhossein Reisizadeh\*, Ramtin Pedarsani, and Salman Avestimehr. “[Coded computing for distributed graph analytics](#),” IEEE Transactions on Information Theory, 2020.
- Amirhossein Reisizadeh, Saurav Prakash, Ramtin Pedarsani, and Salman Avestimehr. “[Coded computation over heterogeneous clusters](#),” IEEE Transactions on Information Theory, 2019.

### Conference and Workshop Proceedings

- Saurav Prakash\*, Amirhossein Reisizadeh\*, Ramtin Pedarsani, and Salman Avestimehr. “[Hierarchical coded gradient aggregation for learning at the edge](#),” IEEE International Symposium on Information Theory, 2020.
- Sagar Dhakal, Saurav Prakash, Yair Yona, Shilpa Talwar and Nageen Himayat. “[Coded federated learning](#),” IEEE Globecom Workshop on Wireless Edge Intelligence, 2019.
- Souvik Kundu\*, Saurav Prakash\*, Haleh Akrami, Peter Beerel, Keith Chugg. “[A pre-defined sparse kernel based convolution for deep CNNs](#),” IEEE Allerton Conference on Communication, Control, and Computing, 2019.
- Sagar Dhakal\*, Saurav Prakash\*, Yair Yona, Shilpa Talwar and Nageen Himayat. “[Coded computing for distributed machine learning in wireless edge network](#),” IEEE VTC Fall Workshop on Vehicular Information Services for the Internet of Things, 2019.
- Amirhossein Reisizadeh\*, Saurav Prakash\*, Ramtin Pedarsani, and Salman Avestimehr. “[Tree gradient coding](#),” IEEE International Symposium on Information Theory, 2019.
- Saurav Prakash\*, Amirhossein Reisizadeh\*, Ramtin Pedarsani, and Salman Avestimehr. “[Coded computing for distributed graph analytics](#),” IEEE International Symposium on Information Theory, 2018.
- Amirhossein Reisizadeh, Saurav Prakash, Ramtin Pedarsani, and Salman Avestimehr. “[Coded computation over heterogeneous clusters](#),” IEEE International Symposium on Information Theory, 2017.

---

\* denotes equal contribution

## PATENTS

- Saurav Prakash, Sagar Dhakal, Yair Yona, Nageen Himayat and Shilpa Talwar. “[Technologies for distributing iterative computations in heterogeneous computing environments](#),” US Patent App. 16/368,716.
- Saurav Prakash, Sagar Dhakal, Yair Yona, Nageen Himayat and Shilpa Talwar. “[Technologies for distributing gradient descent computation in a heterogeneous multi-access edge computing \(MEC\) networks](#),” US Patent App. 16/235,682.

## AWARDS AND ACHIEVEMENTS

- Qualcomm Innovation Fellowship 2021
- Qualcomm Innovation Fellowship Finalist 2019
- Most Novel Research Project Award, EE-599 (Deep Learning course, USC) 2019
- USC Annenberg PhD Fellowship 2016-2020
- Princeton Gordon Wu PhD Fellowship (Declined)
- Viterbi-India Internship 2015
- Summer Undergraduate Research Grant for Excellence at IIT Kanpur 2014
- Shri Singhasan Singh Scholarship at IIT Kanpur 2015
- Institution of Engineering and Technology (IET) Scholarship 2015

## PROFESSIONAL EXPERIENCE

### Graduate Research Assistant

Sep 2016 – 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 logical 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 bipartite 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

### **Applied Scientist Intern**

*Summer 2021*

Alexa AI, 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

### **Graduate Technical Intern**

*Summers 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 the first coded computing framework, CodedFedL, that injects structured coding redundancy into non-linear federated learning for mitigating stragglers and speeding up the training procedure in heterogeneous multi-access edge computing networks
  - Developed a tractable approach for finding optimal coding redundancy and the number of local data points that a client processes during training for minimizing the deadline time
  - Characterized the leakage in data privacy when clients share their local parity datasets with the server
  - Analyzed the convergence rate and iteration complexity of CodedFedL, by treating CodedFedL as a stochastic gradient descent algorithm
  - In numerical experiments using practical network parameters and benchmark datasets, CodedFedL provided gains of up to  $15\times$  in comparison to benchmark schemes

### **International Visiting Student**

*Summer 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

### **Undergraduate Research Assistant**

*2013-2016*

Multimedia Wireless Networks 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

### **Undergraduate Research Intern**

*Summer 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
  - Analyzed capacity lower bound for a MIMO system obeying Block-Fading law using LMMSE estimator for channel estimation at the receiver
  - Obtained an optimal number of transmit antennas for optimizing the capacity lower bound

SELECTED TALKS	<ul style="list-style-type: none"> <li>• TEE-GPU Cooperative Learning: Privacy and Security Without the Price (with Hanieh Hashemi) <i>Presentation, Enclaved AI/ML Workshop, Private AI Research Institute</i></li> </ul>	Jul 2021
	<ul style="list-style-type: none"> <li>• Federated deep learning: On-device learning for CV and NLP (with Chaoyang He) <i>Finalist Team Presentation, Qualcomm Innovation Fellowship</i></li> </ul>	May 2021
	<ul style="list-style-type: none"> <li>• Trustworthy and Scalable Federated Learning (with Salman Avestimehr, Roushdy Elkordy, Chaoyang He, Jinhyun So) <i>CCF Advanced Disciplines Lecture, Institute of Computing Technology, Chinese Academy of Sciences</i></li> </ul>	Apr 2021
	<ul style="list-style-type: none"> <li>• Coded Computing for Federated Learning at the Edge <i>Presentation, FL Workshop on User Privacy and Data Confidentiality, ICML</i></li> </ul>	Jul 2020
COMMUNITY SERVICE	<b>Invited Journal Reviewer</b>	2017-2021
	<ul style="list-style-type: none"> <li>• IEEE Journal on Selected Areas in Communications</li> <li>• IEEE Transactions on Information Theory</li> <li>• IEEE Journal on Selected Areas in Information Theory</li> <li>• IEEE Transactions on Communications</li> </ul>	
	<b>Invited Conference/Workshop Reviewer</b>	2017-2021
	<ul style="list-style-type: none"> <li>• IEEE International Symposium on Information Theory</li> <li>• IEEE Information Theory Workshop</li> </ul>	
	<b>Mentorship</b>	
	<ul style="list-style-type: none"> <li>• Mentor, Graduate Application Mentorship Program (GradAMP), USC</li> </ul>	2021
	<ul style="list-style-type: none"> <li>• Student Guide, Counselling Service, IIT Kanpur</li> </ul>	2013-2016
	<ul style="list-style-type: none"> <li>• Secretary, Fine Arts Club, IIT Kanpur</li> </ul>	2013-2014
	<b>Others</b>	
	<ul style="list-style-type: none"> <li>• Ambassador Caller, Alumni Contact Program, IIT Kanpur</li> <li>• Secretary, Fine Arts Club, IIT Kanpur</li> </ul>	2013-2015 2013-2014
SELECTED COURSEWORK	<b>Algorithms and Artificial Intelligence</b>	
	<ul style="list-style-type: none"> <li>• Deep Learning</li> <li>• Data Structure and Algorithms</li> <li>• Fundamentals of Computing</li> <li>• Machine Learning for Computer Vision</li> <li>• Artificial Intelligence Programming</li> </ul>	
	<b>Mathematics</b>	
	<ul style="list-style-type: none"> <li>• Real Analysis</li> <li>• Probability and Statistics</li> <li>• Linear Algebra</li> <li>• Complex Analysis</li> <li>• Partial Differential Equations</li> <li>• Random Processes in Engineering</li> </ul>	
	<b>Communication and Signal Processing</b>	
	<ul style="list-style-type: none"> <li>• Error Correcting Codes</li> <li>• Wireless Communications</li> <li>• Information Theory</li> <li>• Topics in Cryptography and Coding</li> <li>• Digital Signal Processing</li> <li>• Digital Communication Networks</li> </ul>	
TECHNICAL SKILLS	Python (including PyTorch and Keras APIs), Amazon Web Services, C, MATLAB	