AkshayJajoo

+1-724-397-0561 • akshayjajoo.research@gmail.com

www.cs.purdue.edu/homes/ajajoo • **https://linkedin.com/in/ajajoo**

Current Work

• Researcher at Nokia Bell Labs in Standards Research Group

Education

• **Ph.D.** in Computer Science at **Purdue University**, West Lafayette.
Advisor: **Prof. Y. Charlie Hu.** **Aug' 15 - Oct (Dec)' 20**

• **B.Tech.** in Computer Science & Engineering from **IIT Guwahati**. **Jul'11 - May'15**

President's Gold Medal awardee

. **Dr. Shankar Dayal Sharma Gold Medal** constituted in honour of a past Indian president.

Research Experience & Projects

•Use of spatial dimension for efficient scheduling of distributed jobs, Purdue, USA
Advisor: *Prof. Y.Charlie Hu, Professor, Purdue University* **Sep' 17 - Jan' 20**
Using spatial dimension (number of different machines) for speeding up cloud applications.

•Use of spatial dimension to speed up Coflows and big data jobs, Purdue, USA
Advisor: *Prof. Y.Charlie Hu, Professor, Purdue University* **Sept' 15 - May' 18**
Using spatial dimension (number of different flows) for speeding up communication in cloud applications.

•Summer Internship at Google, Mountainview/Sunnyvale
Team: *NetArch* **May-Aug'17**
I worked with a team which focuses on optimizing and scaling Google's public network's infrastructure. During the internship, along with many advanced Google technologies, I got insight into working of Map-Reduce and distributed jobs by using the internal version of the Google's Cloud Dataflow and Borg.

•Communication System development at TIMA Laboratory, France
Advisor: *Prof. Raoul Velazco, Professor, TIMA Laboratory* **Jun-Aug'15**
Designed and developed a communication system using Arduino processor which provides communication via 3G between one of the experimental FPGA board, and webserver. Also developed the webserver and a user friendly website for managing the experiment.

•Internet Traffic Classification (BTP), IIT Guwahati
Advisor: *Dr. T. Venkatesh, Asst. Professor, Dept. of CSE, IIT Guwahati* **Aug'14-Apr'15**
Identified features to classify HTTP (& HTTPS) traffic from flows using other protocols. LCPSVM was implemented in MATLAB to do this classification. Upto 15% better classification was achieved as compared to normal SVM with these features.

•Poselets for human pose and motion estimation, MMC Lab, Germany
Advisor: *Prof. Rainer Lienhart, Institut für Informatik, Universität Augsburg* **May-Jul'14**
Automated the process of creating random poselets and positive & negative samples for it, implemented and used procrustes distance to compare images closeness. Proposed use of $\beta(2,2)$ distribution for human poselets selection instead of uniform distribution.

•Formulating random algorithm for community detection in graphs, IIT BHU
Advisor: *Prof. K.K. Shukla, Professor, Dept.of CE, IIT BHU* **May-Jul'12**
Worked on developing a random algorithm to find the shortest path between two nodes of a graph and to cluster objects of similar type in the given data.

Publications

•SLearn (A Case for Task Sampling based Learning for Cluster Job Scheduling), USENIX NSDI 2022
Authors: **Akshay Jajoo**, Y. Charlie Hu, Xiaojun Lin, Nan Deng (Google)
The ability to accurately estimate job runtime properties allows a scheduler to effectively schedule jobs. State-of-the-art online cluster job schedulers use history-based learning, which uses past job execution information to estimate the runtime

properties of newly arrived jobs. However, with fast-paced development in cluster technology (in both hardware and software) and changing user inputs, job runtime properties can change over time, which lead to inaccurate predictions.

In this paper, we explore the potential and limitation of real-time learning of job runtime properties, by proactively sampling and scheduling a small fraction of the tasks of each job. Such a task-sampling-based approach exploits the similarity among runtime properties of the tasks of the same job and is inherently immune to changing job behavior. Our analytical and experimental analysis of 3 production traces with different skew and job distribution shows that learning in space can be substantially more accurate. Our simulation and testbed evaluation on Azure of the two learning approaches anchored in a generic job scheduler using 3 production cluster job traces shows that despite its online overhead, learning in space reduces the average Job Completion Time (JCT) by 1.28×, 1.56×, and 1.32× compared to the prior-art history-based predictor. Finally, we show how sampling-based learning can be extended to schedule DAG jobs and achieve similar speedups over the prior-art history-based predictor.

• ***A Case for Sampling-Based Learning Techniques in Coflow Scheduling [Under Journal Submission]***

• ***Philae (Your Coflow has Many Flows: Sampling them for Fun and Speed), USENIX ATC 2019***

Authors: **Akshay Jajoo**, Y. Charlie Hu, Xiaojun Lin

Coflow is a networking abstraction. A group of flows sharing a common end goal are termed as Coflow. State-of-the-art online coflow schedulers in essence approximate the classic Shortest-Job-First (SJF) scheduling by learning the coflow size online. In particular, they use multiple priority queues to sieve long coflows from short coflows, and to schedule short coflows with high priorities. Such a mechanism pays high overhead in learning the coflow size. We propose Philae, a new online coflow scheduler that exploits the spatial dimension of coflows, i.e., a coflow has many flows, to drastically reduce the overhead of coflow size learning. Philae pre-schedules sampled flows of each coflow and uses their sizes to estimate the average flow size of the coflow.

• ***Saath: Speeding up CoFlows by Exploiting the Spatial Dimension, ACM CoNEXT 2017***

Authors: **Akshay Jajoo**, Rohan Gandhi, Y. Charlie Hu, Cheng-Kok Koh

Prior Coflow schedulers approximate the classic online Shortest-Job-First (SJF) scheduling by using a global scheduler to sort Coflows into multiple priority queues, and a local scheduler at each network port to schedule the local flows of Coflows in each priority queue using FIFO. Such a division of the scheduling suffers two problems: (1) The flows of a Coflow may suffer the out-of-sync problem – they may be scheduled at different times and become drifting apart, negatively affecting the Coflow completion time (CCT); (2) FIFO scheduling of flows at each port bears no notion of SJF, leading to suboptimal CCT. Saath is an online Coflow scheduler that overcomes the above drawbacks by explicitly exploiting the spatial dimension of Coflows.

• ***Graviton: Twisting space and time to speedup CoFlows, USENIX HotCloud 2016***

Authors: **Akshay Jajoo**, Rohan Gandhi, Y. Charlie Hu

Graviton provides proof of the concept that how by taking spatial dimension of Coflows into account we can improve Coflow completion time significantly. We have shown that width, number of different ports involved, is a very strong and significant indicator of optimal Coflow ordering.

Courses

CS535	Computer Networks	CS502	Compilers
CS528	Network Security	CS580	Algorithms
CS505	Distributed Systems	CS503	Operating System
CS590	Mathematical Toolkit	CS584	Theory of Computation
EC673	Distributed Computing Systems	CS526	Information Security

Awards and Recognitions

- Awarded with **Dr. Shankar Dayal Sharma Gold Medal** at IIT Guwahati.

This medal is awarded to a graduating student adjudged to be the best in terms of general proficiency including character, conduct and excellence in academic performance, extra-curricular activities and social service.

- Recipient of prestigious **Charpak Research scholarship** for a research internship in France.
- Recipient of prestigious **DAAD-WISE scholarship** for a research internship in Germany.
- Finalist Honda YES award 2013 - 2014.
- Tuition waiver for 4 consecutive years to pursue B.Tech at IIT.
- ISSA member since 2018.
- ACM member since 2017.
- CERIAS member since 2017.

Technical Skills

Programming Languages:	Python, Java, C++, C, Javascript
Operating Systems:	Ubuntu, Windows
Other Technical Skills:	SQL, GIS, Android
Application Softwares:	VIM editor, MS Office, ArcGIS, Android Studio, Eclipse, LaTeX