Raajay Viswanathan

Department of Computer Sciences
1210 W. Dayton Street
Madison, WI - 53706

Phone: +1 (919) 744 9256
e-mail: raajay@cs.wisc.edu
Webpage: pages.cs.wisc.edu/∼raajay

Summary

I am a second year PhD student interested in Computer networks and Machine learning. Prior to grad school, I had worked extensively in data mining and machine learning; applying them to problems in databases and online advertising. Currently, I am excited to tackle problems in computer networks, that involve measurement, data analysis and inference. I believe that software defined networking (SDN) is a convenient platform for addressing such problems; thus building robust, scalable SDN is of interest to me.

Work Experience

Research Assistant, CS. Dept. — Aug 2013 - Present
University of Wisconsin - Madison — Advisor: Prof. Aditya Akella

- SDNTrek: Transition enterprise networks to software defined networks
- OpenNF: A framework for state management in distributed network functions; submitted to SIGCOMM

Research Intern, Machine Learning Group — June 2013 - Aug 2013
Microsoft Research, Redmond — Mentor: Dr. Misha Bilenko

- Implemented a feature exploration algorithm on a new distributed ML framework for *p-click* prediction. Extended the TLC code base.

Research Assistant, Machine Learning & Optimization — January 2011 - August 2012
Microsoft Research India, Bangalore — Mentors: Dr. Srivatsan Laxman & Dr. Prateek Jain

- Cardinality estimation of range queries on relational data, through sparse multi-dimensional histogram estimation from query logs. Tech report at: http://arxiv.org/abs/1111.7295
- Compressed sensing algorithm for multi-label classification of images. Published in NIPS 2012
- Microsoft AdCenter collaboration on smart pricing — proposed an optimization method of fixing discounts resulting in a win-win situation for publishers and advertisers. Successful technology transfer.

Software Engineer, **Cisco Systems Pvt. Ltd.** — August 2009 - December 2010

- Developer in L2 multicast and VPN verticals for Cisco carrier ethernet switches
- Recipient of Cisco star award for excellent performance in the first year of employment

Teaching Assistant, CS Dept., University of Wisconsin-Madison — August 2012 - May 2013

- Lab TA for the *Introduction to Programming in Java* course.

Education

University of Wisconsin-Madison — August 2012 - Present
PhD Computer Sciences Department **CGPA**: 3.857 / 4.0 — Advisor: Prof. Aditya Akella

Indian Institute of Science — August 2007 - July 2009
M.E. Systems Science and Automation **CGPA**: 7.0 / 8.0 — First Class with distinction
Electrical Engineering & Computer Science Department — Advisor: Prof. P. S. Sastry

College of Engineering Guindy, Anna University — July 2003 - May 2007
B.E. Electronics and Communication **CGPA**: 8.6 / 10.0 — First Class with Distinction

Publications

- *Multilabel Classification Using Bayesian Compressed Sensing*
  with Ashish Kapoor, Prateek Jain. NIPS 2012.
- *Discovering Injective Episodes with General Partial Orders*
  with Avinash Achar, Srivatsan Laxman and P.S. Sastry. Journal of DMKD, 2012

Relevant Courses

Advanced Computer Networks, Challenges in Cloud computing, Advanced Machine Learning, Non-linear Optimization, Database Systems, Compiler Optimization, Data Mining, Stochastic Models

Software Skills

- Computer Networking: Floodlight, Pox, Wireshark, GNS3
- Languages & Tools: Python, MATLAB, C++, Java, SQL, SVN, Git, LaTeX
- Experience in Map-Reduce framework — Microsoft's Cosmos & Hadoop

Academic Achievements

- Secured All India Rank 23 out of 30,000 candidates in GATE EC 2007
- Secured State Rank 3 out of 1,50,000 candidates in TNPCEE 2003

Research Project Abstracts

SDNTrek: Exploring Strategies for Transitioning to a Software Defined Network
@ University of Wisconsin Madison, with Aaron Gember and Aditya Akella
Software defined networks (SDN) offer great flexibility in traffic control. However, enterprise network operators are reluctant to migrate to SDN because of the difficulty in transferring existing policies and functions unchanged. Prior work has tackled similar policy management problems by parsing configuration files and modeling routing protocols, but such methods are difficult to scale to large networks and networks with heterogeneous (different syntax for configuration files) devices. We introduce a data analysis framework, to learn existing network policies by feeding packet traces and network logs to a machine learning algorithm. Our method is agnostic to configuration syntax and has the ability to model several classes of network policies (for e.g., end-to-end reachability, rate control). Because of the probabilistic nature of the approach, we cannot guarantee 100% that our method has learnt all the policies (e.g. all ACL rules), since there could be cases that are not presented to the learning algorithm. However, for applications like rate control and traffic engineering that inherently do not demand 100% accuracy, our approach is a generic way of capturing policies that are otherwise hard to learn by existing methods.

OpenNF: Enabling Innovation in Network Function Control December 2013 - February 2014
@ University of Wisconsin Madison, with Aaron Gember, Chaitan Prakash and Aditya Akella
The benefits of using software network functions (NFs), also called middleboxes, and software-defined networking (SDN) together is greater than the sum of its parts. Operators can deploy novel management applications to optimally meet performance, security, availability, and cost objectives for their networks. However, such applications need the ability to effectively manage the processing happening at multiple NF instances in a deployment. We argue that this requires a framework that provides precise, joint control over both the traffic forwarded to, and the internal state maintained at, each NF instance. To this end, we design a control plane called OpenNF. Our design faces several challenges: how to allow applications sufficient flexibility in the objectives they optimize, how to accommodate several types of NFs, and how to avoid possible race conditions when exercising joint control. We address these through careful design of APIs, and two novel constructs—an event abstraction, and two-phase forwarding update. Our evaluation shows that OpenNF offers generally efficient state control without compromising application flexibility, and requires modest additions to NFs with minimal impact on NF performance.

Multilabel Classification using Compressed Sensing February 2012 - June 2012
@ Microsoft Research India, with Prateek Jain and Ashish Kapoor
We present a Bayesian framework for multilabel classification using compressed sensing. The key idea in compressed sensing for multilabel classification is to first project the label vector to a lower dimensional space using a random transformation and then learn regression functions over these projections. Our approach considers both of these components in a single probabilistic model, thereby jointly optimizing over compression as well as recovery tasks. We also derive an efficient variational inference scheme that provides joint posterior distribution over all the unobserved labels. The two key benefits of the model are that a) it naturally allows for handling data that has missing labels and b) provides uncertainty estimates. The uncertainty estimate provided by the model naturally allows for active learning paradigms where an oracle provides information about labels that promise to be maximally informative for the prediction task. Our experiments show significant boost over prior methods in prediction performance with both fully labeled and missing data.

Estimating Self-Tuning histograms from Query Workloads February 2011 - February 2012
@ Microsoft Research India, with Prateek Jain, Srivatsan Laxman and Arvind Arasu.
We propose a general learning theoretic formulation for estimating self-tuning histograms. We cast the problem as an Empirical loss minimization problem by viewing query workload as "training data". We provide formal estimation error guarantees for the class of equi-width histograms. We then go beyond equi-width histograms and present a novel learning algorithm for estimating general histograms. Here, we use Haar wavelets to reduce the problem of learning histograms to a sparse vector recovery problem. Through extensive experiments, we find that our method performs order-of-magnitude better than the current state of the art method for self-tuning histograms.

Finding Frequent Partial Orders from Event Sequences August 2008 - July 2009
@ Indian Institute of Science, with Avinash Achar, Srivatsan Laxman and P.S.Sastry
We present a first-of-its-kind algorithm that could find general episodes with unrestricted partial orders from event sequences. Partial order mining is a much harder problem than finding serial or parallel episodes because of huge number of patterns possible. We point out that frequency alone is not a sufficient measure of "interestingness" in the context of partial order mining. We propose a new measure of "interestingness" for episodes with unrestricted partial orders which, when used along with frequency, results in an efficient scheme of data mining.

References

1. Prof. Aditya Akella,
   University of Wisconsin - Madison,
   e-mail: akella@cs.wisc.edu
2. Dr. Prateek Jain,
   Researcher, Microsoft Research India,
   e-mail: prajain@microsoft.com
3. Dr. Srivatsan Laxman,
   Co-founder, Scibler, Bangalore (prev. in MSR India)
   e-mail: srivatsan.laxman@gmail.com
4. Dr. Ashish Kapoor,
   Researcher, Microsoft Research Redmond,
   e-mail: akapoor@microsoft.com