Sunil Gandhi

Baltimore, MD, U.S.A.
✉ sunilga1@umbc.edu http://www.csee.umbc.edu/~sunilga1/
https://in.linkedin.com/in/gandhisunilr
https://github.com/sunilgandhi007

Education

Current — Ph.D. in Computer Science
University of Maryland Baltimore County
Thesis: "Improving Captioning of video and images"
Advisor: Dr. Tim Oates

May 2015 — Masters of Science in Computer Science
University of Maryland Baltimore County
Thesis: "A Generative Model for Time Series based on Multiple Normal Distributions"
Advisor: Dr. Tim Oates | GPA: 4/4

May 2012 — Bachelor of Engineering in Computer Engineering
Pune Institute of Computer Technology
BE Project: "On demand loading of code in MMUless embedded system"
GPA: 3.3/4

Work Experience

Current / Jan 2013 — Research Assistant at University of Maryland Baltimore County
Cognition, Robotics, and Learning (CORAL) lab
Developed Symbolic Trajectory Analysis and VIsualization System (STAVIS), a system analysis of trajectories using grammar induction algorithm and SAX representation.Also worked on project for noticing changes in the utility of learned knowledge as detecting changes in streaming data. This will be helpful in monitoring performance of neural machine learning algorithms

Jun-Oct 2014 — Research Intern at Xerox Research Center Europe, France
Worked on inferring templates in documents using grammar induction algorithms. We used document classification as evaluation metric for usefulness of grammar learnt from documents.

Publications

- Abhay Kashyap, Lushan Han, Roberto Yus, Jennifer Sleeman, Taneeya Satyapanich, **Sunil Gandhi** and Tim Finin, Robust Semantic Text Similarity Using LSA, Machine Learning, and Linguistic Resources, Language Resources and Evaluation, Springer, to appear
- **Sunil Gandhi**. A generative model for time series based on multiple normal distributions. Masters Thesis at University of Maryland Baltimore County. May 2015.
- **Sunil Gandhi**, Tim Oates, Arnold Boedihardjo, Crystal Chen, Jessica Lin, Pavel Senin, Susan Frankenstein, and Xing Wang. A Generative Model For Time Series Discretization Based On Multiple Normal Distributions. In Proceedings of the 8th Workshop on Ph.D. Workshop in Information and Knowledge Management (PIKM '15). ACM, New York, NY, USA, 19-25.

- Senin, Pavel, Jessica Lin, Xing Wang, Tim Oates, **Sunil Gandhi**, Arnold P. Boedihardjo, Crystal Chen, and Susan Frankenstein. Time series anomaly discovery with grammar-based compression. In Proceedings of 18th International Conference on Extending Database Technology (EDBT). Brussels, Belgium. Mar 26, 2015.
- Kashyap, Abhay, Lushan Han, Roberto Yus, Jennifer Sleeman, Taneeya Satyapanich, **Sunil Gandhi**, and Tim Finin. "Meerkat mafia: Multilingual and cross-level semantic textual similarity systems." In Proceedings of the 8th International Workshop on Semantic Evaluation, pp. 416-423. Association for Computational Linguistics, 2014.
- Senin, Pavel, Jessica Lin, Xing Wang, Tim Oates, **Sunil Gandhi**, Arnold P. Boedihardjo, Crystal Chen, Susan Frankenstein, and Manfred Lerner. "GrammarViz 2.0: a tool for grammar-based pattern discovery in time series." In Machine Learning and Knowledge Discovery in Databases, pp. 468-472. Springer Berlin Heidelberg, 2014.
- Tim Oates, Arnold Boedihardjo, Jessica Lin, Crystal Chen, Susan Frankenstein, and **Sunil Gandhi**. Motif discovery in spatial trajectories using grammar inference. In Proceedings of ACM International Conference on Information and Knowledge Management (CIKM). San Francisco, CA. Oct 27, 2013.
- **Sunil Gandhi**, Swapnil Khorate, Chetan Pachange, Mandar Vaidya. "On-demand loading of code in MMUless embedded systems". Undergraduate Thesis at Pune Institute of Computer Technology May 2012.

Academic Projects

- **Image Description using Deep Neural Networks**
  We used VGGNet model trained on Imagenet dataset to get vectors representing images. We map these image vectors and word embeddings in same space using recurrent neural network that generates description of image.
- **Automating detection of change in distribution in data streams**
  Used A-distance measure to detect changes in data stream generated by neural networks and showed the effect of change of its parameters. The same can be used to monitor performance of machine learning algorithms.
- **Gossip based aggregation on a distributed system using Erlang**
  Created distributed system and calculated max, min, average and median of data at each node using gossip protocol. This system can be used for faster (O(logn)) dissemination of information in a distributed system
- **Using Semantic Technologies to Construct Knowledge base using topic modeling**
  Conference publications, and in turn their authors, can be represented as clusters or well-connected graphs based on their topic commonalities. We created triples representing these associations and performed inference using standard semantic web technologies and ontologies. This will enable a generalized partitioning of graphs based on standard scientific taxonomies such as the Computing Classification System (ACMCCS) and Mathematics Subject Classification (AMSMSC).
- **Analyzing Topic Models Using Separability Assumption**
  We In this project tried to compare several existing topic models and tried to prove empirically that topic model based on Non Negative Matrix factorization generated by making seperability assumption is superior and we try to improve speed of algorithm using random projection
- **On demand loading of code in MMUless embedded system**
  Created the software to modify object code so as to load code in RAM on demand on processors which lacks MMU support. This enables us to load code larger than RAM to be executed in small embedded systems.