SATHAPPAN MUTHIAH

E-mail:sathap1@vt.edu
Arlington, VA | Work Permit: F1-OPT | Phone:801-201-0619
Github: https://github.com/sathappanspm | Website: http://people.cs.vt.edu/sathap1/

SUMMARY

Highly motivated individual with interest in building a career in R&D. Currently, a Phd candidate working in applied Natural Language Processing (NLP) with vast experience in Agile development of AI/ML models. Research focus is in building Hybrid AI systems that can learn to seek human supervision when it doesn't know. Broadly interested in Uncertainty Quantification (UQ), Deep Learning, Transformers, Natural Language Processing (NLP), Information Extraction, Forecasting and Knowledge Graphs (KG).

Projected to graduate in December 2020.

KEY SKILLS

- **Programming:** Python, C/C++, R, Java, Matlab, HTML, PHP
- **Frameworks and Tools:** MongoDB, ElasticSearch, Tensorflow, PyTorch, Scikit-Learn, Numpy, Pandas, Spacy
- **Soft Skills:** Inter-disciplinary collaboration, Agile development, Cross-functional project planning

PROFESSIONAL EXPERIENCE

- **Graduate Research Assistant** — Spring 2012 — Present
  Discovery Analytics Center, Virginia Tech — *Arlington, VA*
  - **Human-aware Approaches for Event Encoding:**
    - Built Hybrid Human-AI system for encoding political events from news media
    - Developed ML system to characterize uncertainty for deciding which documents and tasks require supervision
    - Showcased performance improvement of 14% in Precision, 13.7% in Recall and 12.5% in Quality Score
  - **EMBERS (IARPA Open Source Indicators program):**
    - Worked as a part of multi-disciplinary multi-university team on building a real-time forecasting system of spatio-temporal events like Civil Unrest
    - Led development of model to identify planned events from social media, news and other sources and approaches to create automated narrative (sequence of) stories from real-time news
    - Collaborated in automated geo-coding and model fusion research. Also contributed to overall system architecture and ingestion/enrichment of data sources
    - EMBERS was sponsored by a three-year contract for approx. $13.36M from the IARPA OSI Program
  - **IGACAT (IARPA Functional Genomic and Computational Assessment of Threats Program):**
    - Built DL model based on (CNN, LSTM) and Attention for performing few shot learning to identify if a given nucleotide sequence is a toxin. Work published as poster paper at Machine Learning and Computation Biology conference 2019 (MLCB) and as abstract at ASM BioThreats
    - Researching on Heterogenous interaction based knowledge graphs (KG) for multiple tasks such as identification of genetic functions, possibility of threat, taxonomy etc
  - **Time Series Forecasting for Cyber-Physical Systems:**
    - Investigated utility of Sequence-to-Sequence (Seq2Seq) models for multi-variate time series state forecasting
    - Collaboratively developed Seq2Seq architecture suitable for time-series forecasting and showed 18.83% improvement in load forecasting. Published in IJCAI'16
  - **SAFE (IARPA Mercury):**
    - Directed research efforts for development of fusion system responsible for accepting alerts (tuple of <forecast date,location,actors,type>) from all underlying models and deciding which alerts to merge or suppress
    - Built models to perform real-time clustering of alerts and to predict expected quality of an incoming alert which can then be used to merge/suppress alert
- **Intern, advised by Dr. Prashanta Panigrahi** — Nov 2008 — Jan 2009
  Indian Institute of Scientific Education and Research — *Kolkata, India*
  - Studied the feasibility of a quantum computing system using common and attainable finite dimensional multipartite quantum states
- **Intern, mentored by Dr. Biswajit Chakraborty** — May-June 2010
  National Institute of Oceanography — *Goa, India*
  - Estimated geophysical parameters using SONAR Backscatter data obtained off an experiment on the Western Continental shelf of India. Developed processing and cleansing techniques to process back-scatter data and used simplex method for estimation

SELECTED PUBLICATIONS

- **S. Muthiah**, D. Datta, M. I. Raihan, P. Butler, N. Ramakrishnan, and A. Warren. ProtTox: Toxin Identification from Protein Sequences. In *Machine Learning and Computational Biology (MLCB)*, 2019
- **S. Muthiah**, B. Huang, J. Arredondo, et al. Planned Protest Modeling in News and Social Media. In *AAAI Conference on Artificial Intelligence, January 25-30, 2015*, pages 3920–3927, 2015
- **S. Muthiah**, B. Huang, J. Arredondo, D. Mares, L. Getoor, G. Katz, and N. Ramakrishnan. Capturing planned protests from open source indicators. *AI Magazine*, 37(2), 2016
- **S. Muthiah**, P. Butler, et al. Embers at 4 years: Experiences operating an open source indicators forecasting system. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16
- N. Muralidhar, **S. Muthiah**, and N. Ramakrishnan. Dyat nets: Dynamic attention networks for state forecasting in cyber-physical systems. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*
- P. Chakraborty, **S. Muthiah**, R. Tandon, and N. Ramakrishnan. Hierarchical Quickest Change Detection via Surrogates. *arXiv preprint arXiv:1603.09739*, 2016
- M. R. Islam, **S. Muthiah**, et al. Nactseer: Predicting user actions in social network using graph augmented neural network. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, CIKM, 2019
- M. R. Islam, **S. Muthiah**, B. Adhikari, B. A. Prakash, and N. Ramakrishnan. Deepdiffuse: Predicting the 'who' and 'when' in cascades. In *2018 IEEE International Conference on Data Mining (ICDM)*, 2018
- M. R. Islam, **S. Muthiah**, and N. Ramakrishnan. RumorSleuth: Joint Detection of Rumor Veracityand User Stance. In *2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2019
- Y. Ning, **S. Muthiah**, H. Rangwala, et al. Modeling precursors for event forecasting via nested multi-instance learning. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16

AWARDS AND ACTIVITIES

- **KDD Student Travel Award (2016)**
- **Virginia Tech Pratt Fellowship(2015)**
- **IAAI Deployed Application Paper Award (2014)**
- **Intellectual Property:** P. Butler, **S. Muthiah**, et al. Alert Generation from Multiple Streams, 2016. VTIP 17-047
- **Program Committee:** ACM SIGKDD 2018-20, WWW 2019, ASONAM 2018
- **Talks and Presentation:** ACM SIGKDD 2015-16, IAAI 2014, VT CS Dept. Virtual town hall'20

EDUCATION

- **PhD** in Computer Science — June 2014 — present
  Advised by Dr Naren Ramakrishnan, Virginia Tech — Arlington, VA
  GPA: 3.80
- **MS.** in Computer Science — Jan 2012 — June 2014
  Advised by Dr Naren Ramakrishnan, Virginia Tech — Arlington, VA
  GPA: 3.82
- **B.Tech. in Information Tech.** — 2007 — 2011
  National Institute of Tech. — Bhopal, India
  GPA: 8.65

OTHER PUBLICATIONS

- D. K. Gupta, **S. Muthiah**, D. Mares, and N. Ramakrishnan. Forecasting Civil Strife: An Emerging Methodology. In *HUSO The Third International Conference on Human and Social Analytics*, 2017
- N. Muralidhar, **S. Muthiah**, K. Nakayama, N. Ramakrishnan, and R. Sharma. Multivariate long-term state forecasting in cyber-physical systems: A sequence to sequence approach. *Big Data*, 2019
- J. Schlachter, A. Ruvinsky, L. A. Reynoso, **S. Muthiah**, and N. Ramakrishnan. Leveraging topic models to develop metrics for evaluating the quality of narrative threads extracted from news stories. *Procedia Manufacturing*, 3:4028–4035, 2015
- Y. Ning, **S. Muthiah**, R. Tandon, and N. Ramakrishnan. Uncovering news-twitter reciprocity via interaction patterns. In *Advances in Social Networks Analysis and Mining (ASONAM), 2015 IEEE/ACM International Conference on*, 2015
- N. Ramakrishnan, P. Butler, **S. Muthiah**, N. Self, et al. 'Beating the news' with EMBERS: Forecasting Civil Unrest using Open Source Indicators. In *International Conference on Knowledge Discovery and Data Mining, KDD, 2014*
- A. Doyle, G. Katz, K. Summers, C. Ackermann, I. Zavorin, Z. Lim, S. Muthiah, P. Butler, N. Self, L. Zhao, et al. Forecasting significant societal events using the embers streaming predictive analytics system. *Big Data*, 2(4):185–195, 2014
- Y. Ning, **S. Muthiah**, N. Ramakrishnan, H. Rangwala, and D. Mares. When do crowds turn violent? uncovering triggers from media. In *International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2018