Deepak Gouda

deepakgouda.github.io
linkedin.com/in/deepakgouda
+1 (404) 948-7772
deepakgouda@gatech.edu

Education

Georgia Institute of Technology
Master of Science, Computer Science — Aug 2022 - May 2024
Advisor : Dr. Cecilia Testart — GPA: 4.0/4.0

Indian Institute of Technology, Guwahati
Bachelor of Technology — Jul 2016 - May 2020
Mathematics and Computing (Top 5%) — GPA: 8.80/10.0

Research Statement

My research focuses on comprehensively understanding and improving the Internet and its components through diverse methodologies such as Data Analysis, Internet measurement and Machine Learning techniques. With a strong emphasis on enhancing Internet safety, I aim to develop novel approaches to mitigate vulnerabilities and contribute to a more secure and accessible online environment.

Experience

Research Intern — **Internet Initiative of Japan**
Tokyo, JP — June 2023 - July, 2023

- Performed multi-dimensional analysis of Resource Public Key Infrastructure (RPKI) adoption among Autonomous Systems (AS) to identify key variables that drive the adoption such as - AS topology, location and organization type
- Evaluated effectiveness of Route Origin Validation (ROV) in promoting the adoption of RPKI among ASes, investigating its impact on network security
- Quantified the influence of RPKI adoption on BGP noise of ASes

Software Developer — **Kivi Capital LLC**
Gurgaon, IN — June 2020 - May 2022 (May-July, 2019)

- Led the design and development of low latency algorithmic trading infrastructure, which captured upto 500 GB data per day and executed over 1000 trades/day
- Initiated monthly security assessments to improve the trading infrastructure security
- Maintained a portfolio of trading strategies garnering **100%** returns in **4 months**

Selected Projects

RPKIMetrics - Analysis of RPKI adoption across Internet: *(2023)*

- Goal - identify at-risk business sectors and infrastructure systems susceptible to BGP origin and sub-prefix hijacks
- Identified several variables influencing the adoption of RPKI such as AS size, provider influence, resource ownership time and business type
- Validated the hypothesis that smaller organizations in ARIN are more susceptible to BGP hijacks; results presented at NANOG 89 [Link]

Prefix2Org - Tracing the owners of Internet: *(2023)*

- Developed a scalable mapping of IP prefixes to owners using WHOIS records
- Performed text lemmatization, entity resolution, & fuzzy-string matching to extract useful information from WHOIS text fields and cluster organization information
- Used Prefix2Org for several studies, such as prefix classification by business affiliations and profiling RPKI adoption by BGP prefix origin policies of organizations
- The dataset covers **93.5%** (**95.3%**) of IPv4 (IPv6) prefixes

Achievements

Student Travel Grant — *ACM IMC, 2023*
Student Travel Grant (One of 8 students) — *ACM HotNets, 2022*
Founding member of InfoSecIITG, ranked **298** globally — *IIT Guwahati, 2020*
1st runner in coding hackathon — *Microsoft CodeFunDo++, 2018*
Top **20** among 2700+ teams in Data Science Hackathon — *American Express, 2018*
1st position in Robotics Exposition, Dept. of Defence, India — *DRDO, 2017*

99.2 percentile from 200,000 applicants — *JEE Advanced, 2016*
99.99 percentile from 1.2 million applicants — *JEE Mains, 2016*

Services

Teaching Assistant, Advanced Computer Networking (Spring 2024)
Teaching Assistant, Securing Internet Infrastructure (Fall 2023)
Artifact Evaluation Committee, *ACM CoNEXT 2023*
ACM Professional Member (2023)
Volunteer, EECS Rising Stars (2023)
Mentor, Placement Committee, IIT, Guwahati (2020)
National Service Scheme, IIT, Guwahati (2017-2020)

Publications

- **Deepak Gouda**, Romain Fontugne, and Cecilia Testart, "*RPKIMetrics: Analysis of RPKI adoption across Internet*" [Manuscript]
- **Deepak Gouda**, Alberto Dainotti, and Cecilia Testart, "*Prefix2Org: Mapping BGP Prefixes to Organizations*" [Manuscript]
- **Deepak Gouda**, Hassan Naveed, and Salil Kamath, "*Parameter-free version of adaptive gradient methods for strongly-convex functions*" arXiv:2306.06613, 2023
- **D. Gouda**, S. Jolly and K. Kapoor, "*Design and Validation of BlockEval, A Blockchain Simulator*" International Conference on COMmunication Systems & NETworkS (**COMSNETS**), 2021, pp. 281-289
- De, S., Dey, A.K. & **Gouda, D.** "*Construction of Confidence Interval for Stock Prices Predicted by LSTMs*" **Annals of Data Science.** 9, 271–284 (2022)

Keywords & Skills

Languages: C++, Python, Rust, MATLAB, Bash, LaTeX.
Data: PyTorch, Scikit, Numpy, Pandas, Plotly, Google BigQuery, SQL, Postgres
Networks: Systems, Security, TCP/IP, UDP, BGP, DNS, Routing
Miscellaneous: Airflow, Git, Debian, Linux

Key Courses

Computer Networks : Advanced Computer Networks, Internet Data Science, Securing Internet Infrastructure, Data Communication
Machine Learning : Deep Learning for Text, Big Data Systems, Online Convex Optimization, Statistical Machine Learning, Advanced Linear Algebra

Other Projects

Generative Image Summary and Visual Q&A using LLMs [Github]: *(2023)*
- Developed a pipeline to generate highly detailed summaries of visual content and allow users for question and answers using traditional image captioning models
- Used BLIP-2, GIT and an ensemble of LLaMA-2 and Mistral-7B in the architecture
- Generated summaries are upto **7x** longer than captions generated by state-of-the-art captioning models

Parameter-free AdaGrad for Strongly Convex Functions [arXiv]: *(2023)*
- Developed a parameter-free adaptive gradient optimizer for strongly convex losses
- Devised a hierarchical approach to eliminate manual tuning of the learning rate and the prior knowledge of the degree of strong convexity
- The algorithm achieves a sub-linear regret bound of **O(log T)**
- Method can be extended to develop parameter-free versions of RMSProp and Adam

Text Rating Prediction[Report] *(2022)*
- Developed Machine Learning classifier and regressor models to rate books based on text, summary, author, and publisher information
- Integrated and analyzed data from the CMU Book Summary Dataset and the Google Books API to enhance the prediction model's performance and accuracy
- Achieved best results using GradientBoost (F1 score : **0.72**) and Random Forest Regressor (MAE : **0.29**)
- Key Technologies - *TF-IDF, Universal Sentence Encoder, K-Means, Random Forest*