Falaah Arif Khan

https://falaaharifkhan.github.io/research

EDUCATION:

Shiv Nadar University, Greater Noida GPA: 8.87/10
Aug'14-May'18 Bachelor of Technology-Electronics and Communication Engineering (Distinction)
Minor in Mathematics
Thesis: *Behavioral Biometrics and Machine Learning to secure Web Logins*
Advisor: Madhur Deo Upadhayay (Shiv Nadar University), Sajin Kunhambu (Dell EMC)

WORK EXPERIENCE:

Aug'19-Present **Dell EMC,** ***Software Engineer II***, Bangalore

Dell Cloud Identity (DCI):

>Formulated a game theoretic approach to assigning trust scores to login attempts. The login experience is modelled as a two player, non-zero sum, Bayesian game between the client and the security system. The client has two types, namely malicious or genuine, and the security system takes the best response (authenticating, rejecting or asking for a second factor of authentication) accordingly. Currently prototyping this framework for dell.com login page.

>Created an AI agent that monitors customer satisfaction with the various features of the product, using Natural Language Processing on customer feedback surveys. In addition to quantifying the customer satisfaction score, the model also root causes the source of dissatisfaction into the several features of the product. The model's results are used by product owners to identify useful enhancements to the product.

July'18–Aug'19 **Dell EMC,** ***Software Engineer I***, Bangalore

Dell Access and Identity Solution (DAIS):

> Delivered a behavioral biometrics-based security product that classifies browser activity as being done by a human or by a bot. It is built using an auto associative neural network, trained on client biometrics like keystrokes, click patterns and mouse dynamics. Classifications are made based on reconstruction error between the sample and the prediction. The product is live on Dell login pages, globally.

> Implemented a bot detection filter to identify malicious attempts performed from headless devices (without browser activity). This filter and the classification model together have brought the average number of attempts needed to be manually evaluated down by 87.6% from 125k to around 15k and the average number of accounts that need to be manually protected down by 95% from 800 to 40. In a recent credential stuffing attack, the system was able to successfully thwart 99.71% of the malicious requests on 15390 accounts that were attempted to be taken over.

> Formulated a novel framework to design adaptive robust Completely Automated Reverse Turing Tests that learn on the fly and do not require periodic manual design. The framework works on Adaptive Mechanism design, by learning preferences of human users and custom deep adversaries. The optimal experimental design is learned using Bayesian Inference. The product is live on Dell login pages, globally.

> Created a product traffic forecaster using a Recurrent Neural Network trained on activity logs of the product, to automate the continuous evaluation of product traffic. Insights from the forecaster have been used to alert any anomalous behavior. Instances where traffic has exceeded the baseline have helped identify adversarial action. Deteriorating product health has also been flagged using the baseline, when traffic has fallen below the baseline.

> Architected a graph signal processing approach to dynamic threat modelling. Application logs are used to model the product as a weighted directed graph, where vertices are code elements and edges indicate function calls between elements. Unsupervised learning models are used to set edge weights as indicators of vulnerability to a specific attack. Graph filters are then created and nodes that pass through the filter form the vulnerable subgraph. Superimposing all the vulnerable subgraphs with respect to the different attacks gives rise to a threat model, which is dynamic in nature and evolves as the product grows.

Hackathons and Stretch Projects:

> Designed an Event-based search engine, as an enhancement to conventional image searches. Objects are indexed based on their occurrence at events and this allows for a richer search on lesser input attributes. Bipartite graphs are used as the underlying information structure for search optimization and complexity minimization, while propensity scoring models are used to maximize the precision of information retrieval performed on the graph.

> Designed a novel invoicing system that minimizes errors that occur while converting a purchase order into an invoice. The solution removes redundancies in customer numbers and tax exemption certificates using Tree based methods, improves billing address ranking by using Support Vector Machines on Levenshtein similarity of vectorized addresses and improves the price/discount forecasting model using Multilayer Perceptron Regressors. Currently consulting on this project to help the Invoicing team take the model live.

> Created a robust credit line allocator for the Credit and Collections team, that leverages transactional information along with company data to minimize short term as well as long term risks from overly aggressive and/or conservative credit estimates. The solution models Credit Analyst behavior by training a neural net on historic allocations of credit limits. It also introduces new risk metrics for pay behavior, capacity and credit utilization from unsupervised clustering models. Currently consulting on this project to help the Credit team take the model live.

Jan'18– **Dell EMC, *Intern,*** Hyderabad
April'18 Prototyped a human vs human, genuine/imposter classification model, to identify account match attacks on web logins. Also evaluated the efficacy of this model for a password-less authentication system. Designed a custom dataset of client behavioral information, specifically; mouse dynamics, keystrokes and click patterns. Created a sample dataset from interns typing on the login page hosted in a test environment. Designed the classifier as an ensemble of classifiers, on hand crafted features from raw client data. The model can differentiate two human subjects that are using the same credentials to log in, on the basis of their login behavior.

Aug'17– **Shiv Nadar University, Dept of Electrical Engineering,** Greater Noida
Dec'17 ***Research-cum-Teaching Assistant, Basics of Electrical and Electronic Circuits*** Instructor: Prof RN Biswas (IIT Kanpur)
Designed, conducted and graded weekly lab experiments and tutorials for a class of 30-40 students. Also assisted in conducting and grading the final lab examination.

Aug'17– **Shiv Nadar University, Dept of Electrical Engineering,** Greater Noida
Dec'17 ***Research-cum-Teaching Assistant, Digital Electronics*** Instructor: Prof Sonal Singhal (Shiv Nadar University)
Designed, conducted and graded weekly lab experiments and the final lab examination for a class of 30-40 students.

July'17– **Defence Research and Development Organization, *Intern***, Hyderabad
Aug '17 Supervisor: Dr Laxman Prasad (Programme Air Defence, DRDO)
Studied and simulated security protocols in communication networks.

Jan'17– **Shiv Nadar University, Learning and Academic Support Center,** Greater Noida
May'17 ***Student Tutor, Mathematical Methods*** Course Instructor: Prof Ajit Kumar (PhD, University of Houston)
Curated supplementary course content, delivered lectures and conducted tutorials and practice tests for class of 20-30 students.

PATENTS:

1. Arif Khan, Falaah, Kunhambu, Sajin and Chakravarthy G, K. Behavioral Biometrics and Machine Learning to secure Website Logins. US Patent 16/257650, filed January 25, 2019
2. Arif Khan, Falaah, Mohammed, Tousif, Gupta, Shubham, Dinh, Hung and Kannapan, Ramu. Event-Based Search Engine, US Patent 16/752775, filed January 27, 2020
3. Arif Khan, Falaah and Sharma, Hari Surender. Framework to Design Completely Automated Reverse Turing Tests, in preparation
4. Arif Khan, Falaah and Penugonda, Sai Shreyashi. Dynamic Threat Modelling using Graph Filters, in preparation
5. Arif Khan, Falaah and Swami, Amit. Bot Detection Filter against Credential Stuffing Attacks from Headless Devices, in preparation

PUBLICATIONS:

1. Arif Khan F., Kunhambu S., G K.C. (2019) Behavioral Biometrics and Machine Learning to Secure Website Logins. In: Security in Computing and Communications. SSCC 2018. Communications in Computer and Information Science, vol 969. Springer, Singapore
2. Arif Khan F, Sharma H. "Framework to Design Completely Automated Reverse Turing tests (CART Framework)", in preparation

AWARDS AND HONORS:

August'19	Game Changer Award from Hemal Shah, SVP and Regional CIO, Dell EMC for outstanding innovation in the security features of the Dell Access and Identity Management (DAIS) Product
July '19	Winner, AI Center of Excellence Hackathon, Dell EMC
June'19	Winner, Shark Tank, Dell EMC
April'19	Winner, SafeHack (Security Hackathon), Dell EMC
March'19	Long Term Investment Award from Dell EMC
March'19	Dell Champion Award from Dell EMC
Dec'18	Winner, Hack.Fin (FinTech Hackathon), Dell EMC
Sept'18	Winner, HackLabs Innovation Rally '18 Hackathon, Dell EMC (Theme: Customer Delight)
Aug'14	Recipient of Category 'A' Scholarship (Complete Waiver of Tuition Fees) at Shiv Nadar University

RESEARCH PROJECTS:

Automated Driver, Supervised by Prof CV Jawahar (IIIT, Hyderabad)

Jan '18 – May '18 Implemented basic operations of automated driving, including detecting traffic signs and traffic signal states, identifying obstacles and the subsequent suitable navigation. The automated driver is created by training a convolutional neural network with navigational images and a clustering algorithm using LIDAR data of a bot moving in a simulated environment using Gazebo simulator and ROS.

Booking Assistant, Supervised by Prof CV Jawahar (IIIT, Hyderabad)

Jan '18 – May '18 Created a chatbot that serves as a booking assistant, implementing two skills of making restaurant reservations and movie bookings. Implemented in Python, the bot distinguishes between intents, extracts required attributes and makes suitable recommendations.

Fingerprint Classifier, Supervised by Prof Madan Gopal (IIT Delhi)

Aug '17 - Dec '17 Performed fingerprint classification on the NIST fingerprint dataset, by using Convolution Neural Networks (CNN), as both a feature extractor and a classifier. The features extracted by the CNN were used to train a Support Vector Machine and their classification accuracies were compared.

Tornado Prediction, Supervised by Prof Madan Gopal (IIT Delhi)

Jan '17 - May '17 Forecasted tornado occurrences in the United States, including their magnitude and source (state), using a Neural Network trained on data of tornado occurrences between 1950 to 2009 across the United States published by the Storm Prediction Center.

TECHNICAL SKILLS:

Tools and Languages Python, C#, dotnet, Java, JavaScript, Verilog, MATLAB, LT Spice. Databases: Mongo, SQL, Redis

Certifications 'Foundations of AI/ML', International Institute of Information Technology, Hyderabad (May '18)

Analytical Skills Classification, Regression, Time Series Analysis, Word Embeddings, Clustering, Neural Networks, Tree Based Models, Support Vector Machines, Bayesian Models, Ensembles, Auto-Encoders, Convolutional Neural Networks, Deep Learning Models, Generative Adversarial Networks, Mathematical Modelling, Graph Theory, Signal Processing on Graphs, Algorithmic Game Theory

Project Skills Pivotal (Extreme Programming), Agile

EXTRA CURRICULAR ACTIVITIES:

Model United Nations Participated as a delegate in 27 MUN conferences at both national and international level and won awards in 25 of these, including 10 "Best Delegate" awards. Served on the Executive Board of 12 Model United Nations Conferences in India. Secretary General of MUN Society of Shiv Nadar University (2015-2017)

Literature, Art Writer of https://thefaladox.wordpress.com/
Creator, Writer, Illustrator of 'Meet AI' (comic)

Social Work Manager at AURA: The Social Work and Community Service Society of SNU (2016-2017)