Statement of Purpose

of George Mathew (CSC Ph.D. applicant Fall-2016)

I, George Mathew am a masters student from North Carolina State University(NCSU) in Raleigh. Research and academia has always fascinated me and after deep introspection, I have realized that my greater ambition lies in research and would want to pursue a PhD in Computer Science.

Currently, I work as a graduate research assistant in Real-world Artificial Intelligence for Software Engineering(RAISE) under the supervision of Dr Tim Menzies. My responsibilities include applying machine learning algorithms on software engineering problems like effort estimation and utilizing optimization techniques to solve multi-objective problems software engineering models in Requirements Engineering and Software Quality Prediction. My experience with the RAISE group inspired me to pursue my career in academia and research.

While pursuing my undergraduate degree, my fascination towards robotics motivated me to choose Electronics as my field of specialization. I was part of the robotics society of my university. We were actively involved in creating path follower robots and other such models. Overtime, I started developing a keen interest towards image processing, pattern recognition and machine learning during my junior year. I worked on graphic projects like digital watermark remover for images and rainfall prediction using techniques like neural networks, support vector machines and decision trees. During my senior year I designed and implemented a digital sphygmomanometer. The project introduced me to challenges involved in integrating hardware with software. Since I was trying to solve a very specific problem of capturing signals at very low amplitudes, a customized piece of software was required to solve the challenges of efficiency and stability. After numerous iterations, I zeroed upon the final version, which was adjudged as the best project among over 100 other projects. I graduated from university capping a department gold medal and a university silver medal for my academic accomplishments.

After graduation, I started working as a Software Engineer (level-3) at Payoda Technologies. I was part of a team involved in design and development of a software load balancer called AppViewX. AppViewX is an application centric infrastructure management tool that provides visibility, analytics, security and automation to the network infrastructure. I implemented the load balancer statistics dashboard for AppViewX which collected device statistics on an hourly basis and aggregated over them. Once the statistics were processed, we were able to run predictive analysis over the data to predict faults. My short stint at Payoda Technologies has given me invaluable practical experience in the field of machine learning , distributed computing and network traffic management. I had the opportunity to interact and work with experienced architects and managers thereby giving me both technical and business understanding required to develop a software product. Within 11 months of my professional career I was also awarded the Rookie of the year 2012 and also promoted by two levels to Software Engineer (level-1).

After working in a service based ecosystem for 15 months, I moved to a product based startup called CrowdChat. CrowdChat is a hashtag based platform that enabled users from different social networks to engage in “meaningful” conversations. I was part of the core team which developed the platform using a JavaScript and redis(NoSQL) stack. This experience introduced me to startup culture and the ability to come up with solutions and address real world problems at a rapid pace. Eventually after two years of industrial experience, I decided to pursue a masters degree to broaden my spectrum of knowledge and experience a research and academic lifestyle.

I have been working with Dr. Menzies for over a year. I am funded by NASA Jet Propulsion Laboratory to build a Software Effort Estimator for their space programs. For this project, I collaborated with Dr Jairus Hihn, a senior systems engineer and faculty member of USC and CalTech. The project gave me great insights on how software systems play an important role in space programs and helped me collaborate with an expert, who could use his expertise to validate my experiments. This allowed me to

explore territories in my research like Delphi based learning techniques. Our work was published at the NASA Cost Symposium 2015 and its extension is accepted for publication in IEEE Aerospace Conference 2016. As an extension to software effort estimation, I have worked with numerous statistical non-parametric techniques to rank the estimators used in the project. This paper is under review for publication in the Empirical Software Engineering journal.

During summer 2015, I interned at Facebook as a Software Development Engineer. This experience gave me a holistic view on how to mission critical projects and helped me enhance my understanding on big data analytics. While at Facebook, I worked on open source tools like Hive and Presto which are developed and maintained by Facebook. I applied machine learning techniques like CART and Random Forests from my research to predict performance of HDFS clusters which aided the team to monitor the cluster. The "hackathon" culture at Facebook gave me a platform to collaborate with other engineers working on different fronts across the organization to implement new ideas. This experience helped me develop my team work and adaptability.

Since fall 2015, I am exploring various optimization techniques as part of my masters thesis. Multi-objective problems have always been hard to optimize and software engineering contains many such use cases. My current research focuses on optimizing "softgoals" in requirements engineering. This problem is challenging since the current literature does not solve this problem sufficiently. User feedback in resolving conflicts amongst soft goals is a one such challenge, which can be addressed my enhancing the existing "i*" framework. I hypothesize that using a stochastic approach to traverse the model over a deterministic one would lead to better solutions. The intuition behind this approach is that, it would allow us to obtain a range of feasible solutions rather than a single optimum one.

Apart from academics, I am involved in a few pet projects. I developed an online bookmark manager "region.io", which allows users to store bookmarks and search through them. This was built on top of Apache Lucene, which is used as an indexer for quick retrieval and seamless experience. The site has garnered over 500 users. I have also built a movie recommender system called "octorater" which suggests movies based on IMDB ratings. The recommender system also incorporates feedback based on a simple Naive Bayes classifier. I also have experience as a freelance web developer.

After my PhD, I see myself as an academician and researcher. I believe best research comes from a combination of industrial and research experience. I hope to use my knowledge to contribute actively to the research community. I strongly feel the training would help me enhance my mentorship skills and excel in academia.