Research Statement

My research interests span **Systems and Networking**, particularly addressing the challenges of preventing centralization due to scalability issues while maintaining system simplicity. A pivotal moment that piqued my interest in research in the domain occurred during a Blockchain course in my Sophomore year, where I delved into real-world challenges posed by consensus algorithms. This exploration acquainted me with the shift of these algorithms from Proof of Work (POW) to Proof of Stake (POS). To further understand these shifts, I pursued impactful research in **Blockchain** to understand novel algorithms that address current vulnerabilities in POS algorithms, enrolled in a graduate-level course in **Distributed Systems** where I developed "Raft with persistent Log Replication", and also started working on a project in **Federated Learning**. This experience has equipped me with the foundational experience necessary to embark on a Ph.D. journey, driven to contribute significantly to the dynamic systems research field, particularly in Distributed and Operating Systems.

In the summer following my Sophomore year, I began my research internship with Dr. Zartash Afzal Uzmi in Blockchain, where I focused on independently reviewing existing literature on advanced POS algorithms and their evolution in Blockchain technology. I was particularly intrigued by multiple protocols' novel approaches to achieve liveness and persistence in distributed ledgers while tackling the "Nothing at stake" and "Grinding" vulnerabilities. With a refined theoretical background, my internship became more focused as we developed an equitable "Reputation System" on the Ethereum Blockchain. This was necessitated by the realization that existing systems, particularly in e-commerce, lacked sufficient security measures against collusion, sybil, and re-entrance attacks. To address this with a fair reputation quantification protocol, I used harmonic means to provide weights to each transaction based on the likelihood of collusion and an unfair rating. Upon launching and testing the protocol over Brownie, our quantification model exhibited significant improvement in resistance against collusion and re-entrance attacks. This hands-on problem has taught me on how to navigate through complex, mathematically intensive research papers and discern potential limitations in proposed solutions. More importantly, this experience has prepared me to independently analyze and quantify system vulnerabilities while researching and implementing mitigation strategies.

Towards the end of my Junior year, I and my project partner took the initiative to work on a research project for our senior year under the guidance of Dr. Zafar Ayyub Qazi and Dr. Ihsan Ayyub Qazi. The objective was to understand the implications of Federated Learning (FL) on low-end Android Devices, specifically the Linux Kernel. Currently, state-of-the-art FL applications prioritize accuracy over inclusivity, fostering bias due to the limited participation of low-end devices in developing countries. Furthermore, this issue has gained more traction with the increasing prevalence of FL in applications like Google's GBoard and those using sensitive medical data, driven by HIPAA's strict policies. However, due to such applications being closed-source, the first part of our project was developing our own open-source FL application, which would be capable of running with similar performance, particularly on older Android versions standard on low-end Android devices. For this, we turned to Flower FL, one of the most commonly used frameworks for FL research, except that it struggled with performant compatibility on older versions of Android. Despite this limitation, I independently took the

initiative to adapt the Flower FL framework to the Android architecture, including a service for capturing comprehensive Linux memory states. I programmed an autonomous and self-recoverable background thread to run a TensorFlow Lite model. This advancement ensured the persistence of the FL model's participation with the server via GRPCs despite network failures and device limitations. Developing an independent framework that ran on heterogeneous devices despite the resource and network constraints, when I presented it to Flower's research team, they were very impressed and integrated my contributions into their source code. Having successfully completed a project from inception to product is a testament to my drive to pursue a Ph.D. in state-of-the-art distributed algorithms and address vulnerabilities of large closed-source FL algorithms.

In local testing that followed, we thoroughly examined the outcomes of experiments utilizing the Federated Average strategy and maintaining a consistent model size. Notably, a strong link emerged between a device's RAM size and the memory pressure signals generated from the Linux Kernel during FL training. Deeper analysis revealed high activation of kernel daemons that not only increases CPU consumption but also heightens the risk of FL process being terminated, thus jeopardizing device inclusivity for model training. Concurrently, when assessing user experience by replicating the conditions in which Google conducts its FL training (e.g., during idle times, while charging etc.), we also observed that other applications (e.g., YouTube, Google Chrome etc) experienced significant performance drawbacks, contradicting Google's claim. This presents another exciting tangent for our project, and we intend to rethink Google's strategies for more inclusivity of low-end Android devices. So far, despite the challenging journey, the project has proven to be immensely fulfilling, honing my ability to question prevailing paradigms and statistically validate their shortcomings by exploring previously uncharted domains.

As a Teaching Assistant for a graduate-level course, Distributed Systems, I've cultivated a passion for teaching. I adeptly address students' challenges, emphasizing foundational understanding. Leading tutorials on complex systems, debugging in Go, and crafting Raft test cases has fortified my ability to create challenging assignments, showcasing my dedication to fostering student comprehension.