Individuals and enterprises produce over 2.5 exabytes (10^18 bytes) of data everyday. Much of this data - including sensitive and private information - is stored with and managed by third parties, such as Amazon Web Services or Google Cloud. These companies can lose millions to billions of dollars in sales if their data access latencies increase by only a few hundred milliseconds [2]. Hence, reducing data access latency to improve performance received the highest priority while designing cloud data management systems. But the ever growing number and sophistication of cyber attacks on the cloud [3] coupled with increases in legal requirements for data privacy and security (e.g., GDPR or HIPAA) [1] have forced cloud providers to re-evaluate their priorities. However, there exists a fundamental trade off between security and efficiency in data management systems. While resolving this tension is challenging, it has fostered the growth of a deep field at the intersection of cryptography and database research. In Springer, for example, the number of published papers in this intersection grew by 50.1% since 2015 and 91.8% since 20001.

My PhD research encompasses designing, prototyping, and evaluating data management protocols that strike a balance between efficiency and security in both trusted and untrusted environments. Before being able to solve security challenges in database systems, I first had to fully understand existing system designs. My extensive study of distributed data management, where data is partitioned and/or replicated across geo-distributed locations, led me to identify open problems in traditional, trusted cloud databases. The systems I built in trusted clouds perform 16x-18x better than the state of the art [6, 9]. As the next step towards building secure data systems, I developed a fault-tolerant data system for hybrid clouds that consist of a small trusted private cloud and a larger untrusted public cloud [5]. This system tolerates malicious failures in the public cloud and crash failures in the private cloud, and by leveraging the locality of crash and malicious failures, our system performs 37.5% better in throughput than the state-of-the-art baseline systems. Finally, to better understand the interplay between various security guarantees and performance, I conducted a tutorial on blockchain systems [11] – an ideal example where untrusted geo-distributed entities manage critical data. Equipped with blockchain techniques that protect data, I built secure database systems that guarantee data integrity [8] and data privacy [7, 10]. I developed the first distributed transaction commitment protocol [8] that tolerates up to n – 1 maliciously failing servers (out of n servers) without using expensive data replication and the first fault tolerant and highly available oblivious datastore [10] – one that hides access patterns of users along with the contents of data.

Moving forward, my research vision is to bring feature-rich databases that are secure and high performing into reality. Many existing oblivious database systems trade useful database features such as concurrency control or query optimizations in favor of performance; this leads to a limited backend database, which supports only simple Get and Put operations. Building feature-rich private databases that are also practically usable requires experience designing database features in trusted settings, and the ability to identify security weaknesses and choose the best cryptographic tool to transform a non-private feature into a private one. While I have already established my ability to develop private and secure database features such as transaction commitment [8] and fault tolerance [10], I am interested in furthering this research area by creating private versions of other features such as concurrency control and query optimization. Having demonstrated a track record of building high performing secure databases, I am the ideal candidate to realize the vision of integrating vital features into these secure systems.

Trusted Infrastructure

Before delving into untrusted settings, I started my research career in traditional cloud settings, which assume trust. Cloud enterprises typically replicate their data to provide fault-tolerance and shard (or partition) their data and store the shards on multiple servers to provide scalability. State of the art databases, such as Google's Spanner, use atomic commit protocols (e.g., Two Phase Commit) to allow scalability and consensus protocols (e.g., Paxos) to achieve replication. Spanner treats atomic commitment and consensus disjointedly, where it hides the consensus logic from commitment logic and vice versa. Motivated by the co-existence of sharding and replication in most real-world databases, our work [9] unifies the two seemingly disparate paradigms into a single framework called Consensus and Commitment (C&C). The C&C framework is a great pedagogical tool as it can model many established data management protocols, while also providing insights to propose new ones. To highlight its advantages, we propose a novel commit protocol, G-PAC, which merges consensus with commitment. As a key feature, unlike many existing protocols that separate failure recovery logic from failure free execution logic, G-PAC executes the same set of instructions for both scenarios, easing a developer's job. The unified approach of G-PAC reduces one (out of three) round of cross-datacenter communication compared with Google's Spanner; this allows G-PAC to perform between 27-88% better than Spanner-like protocols in terms of throughput.

While G-PAC extended the C&C framework to the context of distributed transactions for generalized workloads, we developed Samya [6] to extend the C&C framework to manage high contention, hotspot data. Samya, a geo-distributed data management system, stores and manages hotspot aggregate data (e.g., number of cloud resources or number of items in stock).

---

1I searched for conference papers with keywords database AND security OR privacy OR encryption on https://link.springer.com/search.
State-of-the-art geo-distributed databases such as Google's Spanner take a centralized approach where a server acting as a leader processes all client requests to a specific data item sequentially, which thwarts throughput and leaves the data replicas underutilized. To utilize all replicas or sites that store a data item, Samya splits or partitions the aggregate data, modeled as tokens (e.g., resource tokens), and stores individual partitions on different servers. This design choice allows servers in Samya to independently and concurrently serve client requests. A site serves requests locally as long as it has locally available tokens. Samya employs predictive models to predict if a site will exhaust its tokens; if so, the sites execute a novel synchronization protocol, Avantan, extended from the abstractions of the C&C framework, to redistribute the spare tokens in the system. Compared with the centralized solution, Samya's parallelism reduces the 99th percentile latency by 76% and allows Samya to commit 16x to 18x more transactions.

Hybrid Infrastructure

To accomplish my goal of building trusted databases using untrusted infrastructure, I started examining hybrid cloud infrastructures encompassing a small trusted private cloud and a larger untrusted public cloud – a logical middle-ground between trusted and untrusted settings commonly deployed today. The default protocol choice to replicate data in such hybrid clouds has been PBFT, a Byzantine fault-tolerance protocol, which does not distinguish between crash or malicious failures and is unaware of the locality of the failures. In our work, SeeMoRe [5], we propose replication protocols for hybrid cloud environments that leverage the trust of a private cloud and the scalability of a public cloud to achieve better performance compared with PBFT. SeeMoRe tolerates crash failures in the private cloud and malicious failures in the public cloud. This work highlights how we can develop efficient protocols by being aware of where different types of failures (both crash or malicious) occur in hybrid cloud environments. Specifically, compared with PBFT, SeeMoRe reduces the number of communication phases and messages exchanged and also requires fewer replicas to tolerate a given number of malicious and crash failures, reducing an application's overall deployment cost. We developed SeeMoRe to dynamically switch between three different modes of operation depending upon the load on the private cloud: higher load on the private cloud leads to availing more servers from the public cloud and vice versa. Our extensive evaluation revealed that SeeMoRe performs 37.5% better than PBFT at peak throughput and only 8% worse than the state of the art protocols that tolerate only crash failures.

Untrusted Infrastructure

After working with fully trusted infrastructures and hybrid environments, I became intrigued by the question, "What if I host my data on completely untrusted infrastructure?". To answer this question, we propose Trust-Free Commit (TF-Commit), a trust-free commitment protocol that executes transactions across multiple untrusted servers [8]. To our knowledge, TF-Commit is the first atomic commitment protocol to handle malicious failures without using expensive Byzantine replication. Byzantine replication protocols tolerate less than n/3 malicious failures, where n represents the total number of servers, whereas TF-Commit tolerates up to n−1 malicious failures. TF-Commit combines Two Phase Commit with a collective signature scheme, CoSi, to commit transactions across multiple untrusted servers. Similar to blockchain, committing each transaction (or a set of transactions) produces a tamper-resistant log entry, which each server appends to its log. The tamper-resistance stems from collective signature created during commitment that cannot be modified by a single server. An external auditor then audits this log to detect any faulty behavior. TF-Commit's failure detection mechanism precisely identifies the point in the execution history at which a fault occurred as well as the servers that failed. These guarantees provide two fold benefits: (i) An auditor always detects both a malicious fault and the misbehaving database server, and (ii) A benign server can always defend itself against false accusations. Many scenarios can benefit from TF-Commit, such as blockchain systems or supply-chain management where the participating entities span different administrative domains that do not trust each other. Compared to executing a transaction in trusted infrastructures, executing TF-Commit has only 1.8x overhead in latency and 2.1x in throughput - an acceptable overhead for many use cases given the additional security guarantees of TF-Commit.

While TF-Commit ensured data integrity, the growing number of data leaks in the recent past inspired me to work on data privacy. Recent attacks revealed the inefficiency of mere data encryption to protect data privacy; an attacker can uncover non-trivial information either about the data or its users by observing the users' access patterns. Oblivious RAM, or ORAM, is one of the cryptographic techniques that prevents access pattern attacks. Although the database literature consists of many ORAM-based systems, to-date no ORAM-based datastore supports fault tolerance – an important primitive in database systems. To this end, our work, QuORAM, proposes a quorum based replicated ORAM datastore that tolerates crash failures while preserving obliviousness [10]. QuORAM consists of multiple untrusted and potentially colluding storage servers, each accessed via a separate trusted proxy. QuORAM guarantees linearizable semantics – all operations on a data item appear to be linear – using a lock-free replication protocol where a client always reads from a quorum (e.g., majority) of replicas and writes to the same quorum, for both read and write requests. If proxies treat the reads and writes as separate ORAM requests, then they fetch the path twice sequentially from the storage server, which leads to prohibitive latencies. To avoid double-fetching, proxies in QuORAM maintain an incompleteMap to store request mappings of requests that are read but not yet written back. Our evaluations of QuORAM reveal the advantages of geo-replication in ORAM systems: reduced latency for geo-distributed
clients and reduced load on a single proxy. QuORAM reduces the average data access latency by 54.4% and improves the throughput by 51.2% compared to a non-replicated ORAM system, while providing fault tolerance.

Future research interests

The use of smart devices – wearable health monitors, home assistants, bio-metric security appliances – is on the rise. These smart devices typically communicate with the cloud, either directly or indirectly, to store and retrieve an end user's data. Such outsourcing of data necessitates building trustworthy data management systems where the data can reside atop of potentially untrusted infrastructure. While trusted cloud infrastructures are feature-rich and relatively mature, advanced features in untrusted cloud infrastructure remains a burgeoning technology that places a high burden on application developers and has expensive deployment costs. For my future research, I will continue to design efficient and secure protocols and systems to manage data, and pioneer new research on building cost effective and feature-rich private databases catering to specialized domains such as Internet of Things (IoT) devices. Specifically, I have three focus areas for my future research:

Serverless Paradigms

Serverless computing is a paradigm where the cloud provider dynamically allocates and manages resources, allowing a developer to focus solely on application development. Serverless computing executes computations in bursts without retaining any execution state in memory for long. Many applications find it appealing due to its cost effectiveness and low overhead for the developers. I envision many existing applications will switch to serverless paradigms, and in fact, recent data forecasts the serverless market to grow from $7.6 Billion in 2020 to $21.1 Billion by 2026 [4]. Assuming trusted settings, existing transactional applications cannot be easily migrated due to the inherent stateless nature of serverless architectures. Transactional workloads require maintaining state, typically distributed across servers, and the servers that maintain the state need to coordinate to ensure the state's consistency. While some recent works provide transactional consistency and isolation in serverless settings, the space of ACID-compliant (Atomicity, Consistency, Isolation, and Durability) datastores remains widely unexplored.

I have started exploring ways to incorporate ACID guarantees in a serverless platform. As a research intern at IBM Research, I worked on building a transactional framework on top of a serverless runtime, KAR², developed at IBM. The transactional framework² provides atomicity, strong consistency, serializable isolation, and durability (ACID) guarantees. While analyzing the performance of KAR's transactional framework using the TPC-C benchmark, I realized that serializable and strongly consistent data impose high performance costs, indicating a serverless approach may not be correct for these data types. In my future research, I plan to answer the question: what consistency and isolation guarantees best suit serverless architectures? Specifically, I plan to develop an ACID-compliant transactional data system with snapshot isolation (an isolation guarantee more relaxed than serializability) and eventual consistency (weaker than strong consistency) and compare the performances with their stricter counterparts. Both snapshot isolation and eventual consistency are practical guarantees that database systems such as Amazon's DynamoDB and Microsoft’s CosmosDB provide.

Bridging the Functionality vs. Security Gaps in Oblivious Datastores

Database researchers, and people in computing in general, are grappling with the increasing challenge of data privacy. Laws such as GDPR and HIPAA make data privacy not merely a desired property but a necessary property that database systems must provide. Merely encrypted data is susceptible to privacy attacks based on access patterns, and Oblivious RAM or ORAM, a cryptographic technique, protects data from access pattern attacks. Although database systems have evolved significantly over the years to provide a rich set of features such as concurrency control, transactional ACID guarantees, and query optimizations, almost all currently existing privacy preserving oblivious datastores strip away these features and downgrade the untrusted backend server to a simple system that supports single item Gets and Puts. This simplifying of the backend database servers is usually necessary to uphold the strict definitions of obliviousness, as different database features may reveal non-trivial information about the data. For example: allowing the untrusted database server to handle concurrency control may reveal high contention in application workloads, as multiple concurrent requests might be accessing the same data item. Today, a trusted proxy server provides much of these features such as concurrency control and the backend server merely Gets and Puts individual data items.

I am interested in developing feature-rich databases, where the features themselves preserve privacy, and to remove the trusted proxy, allowing the clients to access the database servers directly. To achieve this, I aim to consider each database feature separately and build oblivious versions of those features. For example, concurrency control can be made oblivious by creating fake accesses to data items and mixing them with real access such that from the server's perspective, all data items have equal concurrency. I believe such a feature-by-feature approach to integrate obliviousness into database systems can help bridge the gaps between theoretical oblivious datastore constructions and existing practical databases.

Fault-tolerance and Security of IoT data

The number of IoT devices in use is speculated to reach 27 billion by 2025, while we use 12.3 billion IoT devices today. Users find IoT devices highly appealing as they allow them to access a myriad of applications and services with a single touch, such as paying with a tap of a smartwatch or phone. Since IoT devices can span

²https://github.com/IBM/kar
³https://github.com/IBM/kar/tree/main/examples/actors-transactions
sensitive domains such as health monitoring or military systems, data produced by certain IoT devices must be fault tolerant and the entire data collection and processing pipeline must be secure, without overly degrading performance.

Edge computing or edge cloud, a relatively new paradigm, enables quick storage and compute access to edge devices such as IoT devices. Typically, edge devices push their data to edge nodes – servers physically close to edge devices – and these nodes then push the data to a persistent and fault-tolerant cloud storage. While the cloud has fault tolerant storage, the edge nodes generally do not. Fault-tolerance techniques such as data replication using Paxos are unfit for edge cloud due to their large communication overheads. As a future research goal, I will develop novel fault tolerance techniques for edge cloud and compare them with state of the art techniques such as replication using Paxos. Specifically, I aim to explore eventually consistent fault tolerance techniques where the main thread of a primary node continues to process client updates and a background thread updates the secondary (and potentially other) node(s).

With regard to privacy and security of IoT data, the literature consists of many works on securing the communication layer (i.e., networks) and solutions to encrypt the IoT data for privacy. In the data processing pipeline of IoT devices, encrypted data eventually migrates to the cloud. As described earlier, mere data encryption is insufficient to achieve data privacy due to access pattern attacks. Presently, no current work has studied the cloud storage access patterns unique to edge devices. Existing solutions to mitigate access pattern attacks, such as ORAM, are impractical for IoT devices due to their prohibitive latencies. In the future, my research group will study the access patterns unique to IoT/edge devices, identify specific attacks plausible due to such patterns, and devise novel solutions to mitigate access pattern attacks while retaining the low access latency necessary for IoT devices.

Research Funding

I understand the importance of raising funds to support my future students and to materialize my future research goals. I showcased my ability to obtain funding for my research by receiving two fellowships to fund my PhD: IBM PhD Fellowship totalling $95,000 over 2 years and Google PhD Fellowship (declined) with net worth $159,000 over 3 years. I also secured funded conference travels to attend Rising Stars (awarded to selected women PhD students and postdoctoral scholars interested in academia), CRA-W Women’s Grad Cohort, Grace Hopper Conference, and a top database conference, VLDB (2018). At UCSB, I contributed to writing a NSF CNS grants that partially funded my PhD. As a faculty, I plan to submit grant applications to NSF CNS, IIS as well as OAC for my projects on data security and privacy. I have also been forging industry connections with IBM, Facebook, and Google, and I plan to apply for industrial faculty research or early career awards.

References


I believe course instructors and teachers can make or break not only a subject, but an entire career path for a student. Anecdotally, my high school Mathematics teacher helped me identify my love for abstraction, which motivated me to pursue Computer Science. My undergraduate Computer Networks professor, through clear course design, sparked my interests in backend programming when I had always been intimidated by it. When starting my graduate career at UC Santa Barbara, I was unsure of pursuing a PhD, in part due to feeling excluded within the first research group I was a part of. Fortunately when I took a course with my current advisor, his evident passion for distributed systems influenced me to pursue research opportunities with him. His encouragement, guidance, and friendship played a critical role in my decision to switch from Master’s to the PhD program. I am fortunate to have the mentors I did, and as a professor I am committed to creating such a positive impact in the lives of my students as well.

Computer Science today does not represent the diversity of the general population; these ratios are more skewed in graduate school and even more at the faculty level. My experience has taught me how significant a connection with a mentor can be in encouraging and retaining students who are a minority within academia in general, and Computer Science in particular. As a professor I will strive to: i). employ inclusive teaching strategies and stay abreast of current pedagogical research both to teach more efficiently and to retain minority students, ii). instill critical thinking skills and promote a growth mindset both within my lab and within the lecture hall.

Teaching Experience

The two major sources of my teaching experience stem from being a Teaching Assistant at UCSB and volunteering as a tutor at School On Wheels, a non-profit that tutors children experiencing temporary homelessness.

At School On Wheels, my goal as a tutor is to develop a growth mindset in my students and I focus on the students’ attitudes more than academic materials. These students, typically experiencing extreme challenges, commonly say ‘I don’t know X [math or spelling words]’ and I always correct them and ask them to instead say ‘I don’t know X yet’. Since many of these students experience attention disorders, I employ creative teaching techniques such as including art with words or jumping jacks with numbers. To counteract the chaos typically prevalent in my students’ lives, my tutoring sessions follow a clear structure, starting with a small mindfulness exercise. I also discuss the benefits of education and the importance of a career with my students by sharing anecdotal experiences or watching motivational videos together. Through this volunteering, I hope to encourage more underprivileged students, many of whom are underrepresented, to join STEM.

At UCSB, my goal as a Teaching Assistant (TA) has been to instill critical thinking in my students. I served as a TA for 7 courses, both within and outside my research field. As a TA, I designed and evaluated programming assignments, quizzes, and final exams in addition to holding weekly discussion sections reinforcing topics covered in the course. But the most exhilarating teaching sessions were when I covered lectures (over 6), including graduate level courses on Distributed Systems, for my advisor. Experiencing the sudden look of understanding in students’ eyes when teaching complex concepts, such as Paxos - a convoluted multi-phase protocol, was a joy worth experiencing in spite of preparation time spent away from my research. To encourage critical thinking, rather than present the concepts as-is to my students, which can be overwhelming, I try to build intuitions by providing naive solutions and ask the students to identify issues in the naive solutions. I will always remember a student’s feedback after I taught Paxos, when they said I helped them understand the protocol better than my advisor, who I consider an exemplary lecturer. I break a single lecture into logical parts and after each block I summarize the part and strongly insist the students to ask questions. After the awkward first 30-60 seconds, I almost always receive questions. By staying in class after the lecture I provide students, shy of asking questions in class, an opportunity to clarify their questions.

Teaching methodology

My teaching methodology will focus on critical thinking and employ inclusive strategies to engage all students, which directly helps retain minority students in Computer Science. As a faculty, it is just as important to keep up with evolving teaching strategies as well as evolving technical research developments. The courses I design will include inclusive teaching strategies backed by recent and evolving research [4, 3]. Specifically, the courses will:

- **Be clearly structured**: I will design courses with a clear structure and communicate the course plan at the beginning so as to minimize catching students by surprise. I already practice defining clear structures to my tutoring sessions at School on Wheels, where my students come with a different mood each session. Similarly, since a class typically consists of students from different learning abilities, confidence levels, and cultural backgrounds, having a clear structured class with known expectations may benefit most students, especially the minority students.

- **Connect CS concepts to real world examples**: I plan to form analogies between complex CS concepts with real world, well known scenarios. For example, in lower division classes, I can design a programming assignment to implement a linked

---


2. [https://schoolonwheels.org/](https://schoolonwheels.org/)
list to read DNA sequence data. Or in upper division lectures, I will use analogies such as consensus protocols (e.g., Paxos) being equivalent to a government passing a bill, or commitment protocols (e.g., Two Phase Commit) to two people getting married. I have employed such examples and analogies during invited talks and have received positive feedback on them.

- **Include projects**: Being a systems person, the courses I teach will include programming assignments, which I believe are vital to bridge the gap between theory and practice. As a TA, I have experience designing programming assignments and forming clear rubrics to test the implementation. When I was a TA for distributed systems course, far more students asked detailed questions on concepts when they implemented Paxos, a complex protocol, than when they had a quiz on Paxos.

- **Consist of low-stake quizzes over large exams**: Research indicates [4] that smaller but multiple low-stake quizzes form inclusive teaching whereas one or two large, high-stake exams serve the opposite. I believe small quizzes or tests makes a class less intimidating and provide more frequent opportunities for students to assess their strengths and weaknesses. I am also a proponent of open book tests: during my undergraduate studies, I learned the most from Computer Networks class, which consisted of open book exams. I believe open book tests evaluate a student’s grasp on the subject, without the side effects of how much a student can memorize concepts or equations.

Further, my teaching strategies will be kept up-to-date by following guidelines such as ACM’s EngageCSedu [2], which consists of evidence-based techniques to engage and retain students, and CS Teaching Tips [1], which includes easy-to-implement tips on fun ways to teach CS concepts and tips to reduce bias while delivering lectures and evaluating students.

### Research Mentoring

At UCSB, I have had the opportunity to closely mentor 9 students in different contexts. In the research context, I mentored 2 undergraduates, 3 master’s, and 2 junior PhD students, which led me to co-author research papers with 6 of my mentees. As an international student, I struggled with clearly articulated academic writing in my PhD. When reaching out for advice, I learned about an online scientific writing course that helped me tremendously. As a mentor to junior PhD students in our lab, I encouraged my mentees to enroll in the online course early on in their PhD and as a result, we all saw a noticeable improvement in their writing quality. With the undergraduate and master’s degree students, I mentored a diverse set of students: some required hands-on mentoring while others preferred hands-off; some were highly motivated and delivered results early, while some others needed a nudge. With one student in particular, I found that I had to probe at every step of a project and knew this could not be a permanent solution. To solve this, I shifted from asking about the specific results I was expecting to be done to asking the student what their next few days consisted of and when they thought they could deliver the next step to me. This seemingly minor shift led to the student being more comfortable in sharing some of the challenges they were facing and also made them feel more accountable towards the project.

In my academic career, I also had the opportunity to mentor 2 junior female graduate students, where I met with my mentees every week for their first quarter at UCSB. My goal for this mentoring was to provide early support to incoming students by helping them address any of their questions or concerns. When one of my mentees indicated her interest in business management, I connected her to an alumni with a business management degree from UCSB, which enabled my mentee to receive dual master’s degrees. I helped my other mentee, who had qualms about her research lab, identify and switch labs that fit her interests.

I believe my research and general mentoring experiences so far have prepared me well to take on the role of an assistant professor. As a faculty in charge of my lab, I will adapt my mentoring strategies to cater to an individual student’s preferences, strengths, confidence, and prior experience. As a minority in Computer Science and having reached this career stage with roots from a small town in India, I understand that each of my students will come with a different backstory and that there is no one mentoring strategy that fits all. Apart from PhD students, I intend to mentor master’s and undergraduate students, and even high school students. In recruiting students, I will look for students with curiosity and a growth mindset over an academic success record. Further, I will learn about all the mental health help that the university or the department provides and by openly communicating about the mental health help I benefited from in my PhD, I will advocate my students to seek help if and when needed.

### Summary

Enthusiastic and passionate teachers helped me shape my career; understanding the significance of a teacher first hand, I vow to share my enthusiasm and passion for Computer Science when I teach courses. I am interested in teaching undergraduate level courses on Introduction to Distributed Systems, Database Systems, Computer Networks, Operating Systems, Data Structures and Algorithms, and other lower level courses. At the graduate level, I am looking forward to teaching Advanced Distributed Systems, Advanced Databases Systems, Designing Encrypted and Oblivious Databases, Permissioned and Permissionless Blockchains, and seminars or reading groups on these topics. Given my teaching experience and my commitment to inclusive teaching philosophy, I believe I am a good fit to be a Computer Science faculty. I would love to see more women faculty in Computer Science and engineering. By applying to this role, I am trying to be the change I want to see in the world.
References

Diversity, Equity, and Inclusion Efforts

Efforts within Computer Science

As a PhD student, I have strived to enhance diversity not only within the Computer Science department at UCSB but also within the database field, my research area. I served as an elected Graduate Student Representative (grad-rep) for 3 years at UCSB. Recently, as a grad-rep for faculty recruitment, my interests aligned with the department's mission to increase diversity among faculty (less than 14% faculty were women then). As a grad-rep I contacted Rising Star award recipients, an award that recognizes women graduate students and postdocs interested in academic careers. Further, I considered students’ feedback valuable in deciding which candidates to hire, especially because graduate student population had higher diversity (28% female) than faculty (14% female faculty). Thus, I actively sought students’ feedback on the candidates and conveyed it to the hiring committee, urging them to consider our feedback. All our efforts resulted in 43% of the short-listed applicants being women, whereas the previous year was 22%, and one of them received an offer (out of 3 offers sent) and is now a faculty member at UCSB. Apart from serving in faculty hiring committee, I also served as the Talks&Events representative to strengthen the social ties between graduate students and the faculty. I organised weekly and monthly coffee hours and social hours. The coffee hours successfully gathered students across different labs and the technical discussions there led to many collaborations across labs. While minority students have some unique challenges, all graduate students have the challenge of constant stress, and events like coffee hours help alleviate the stress by forging friendships, creating inclusive atmosphere, and pollinating new research ideas within the department. The department recognized my diligent efforts towards its welfare and in 2019 I received the Outstanding Graduate Student award in the department – making me the youngest student to receive the award.

Within the database community, I am a core member of the DB D&I ¹ (Diversity and Inclusion in Databases) initiative. As the only graduate student core member (out of 10 members), I co-lead the SCOUT initiative wherein we identify meaningful D&I efforts in other CS communities and advocate for those efforts within the DB community. I am also in-charge of co-writing the Code of Ethics that can serve as a guideline for all future database conferences. I served as a D&I session panelist at SIGMOD 2021², a top-tier DB conference, voicing the international students’ challenges in attending in-person conferences and the challenges we faced during the pandemic. My service as a grad-rep and my early involvement in DB D&I efforts indicate that I can successfully take-up leadership roles and channel my abilities towards increasing the diversity in Computer Science.

Outreach beyond Computer Science

‘Opportunities have to be equal before you can know if abilities are equal’, Melinda Gates. Equitable opportunity to people of all genders, races, and socioeconomic classes is an important cause to me. Education can be an equalizer and help disadvantaged people with socioeconomic upward mobility [4]. Anecdotally, none of my grandparents, farmers from rural India, finished middle school; starting at age 8, my father waited tables during the day and attended school and college in the evenings to support himself and receive a bachelor’s degree, and my mother was the first woman from her village to attend college. Their education uplifted our lifestyle and as a direct result, I am on track to receive a doctorate degree. I truly believe that education helps with upward mobility, and to this end, I volunteer at School on Wheels³ – a non-profit dedicated exclusively to the educational needs of children from k-12 experiencing homelessness. I tutor students ranging from six year olds to late teens and in addition to academic help I act as a mentor, positively reinforcing their attempts at solving problems. My current 1st grade student struggles with attention disorder and oftentimes I have to devise novel ways to engage her where we color pictures, say of apples, before learning to spell apple or do additions followed by that many jumping jacks together. I cannot articulate the joy she brings me when she reads or writes a long, complicated word! With her and my other students, more than anything, I aim to inculcate a growth mindset and encourage them to believe in themselves. With this volunteer effort, I hope to convey the importance of

education and convince my younger students to finish high school and my older students to consider college, when they often never had before.

**Personal Challenges and How They Influence My Future Goals**

Being a woman, a woman of color, and an immigrant (or rather, a ‘non-immigrant alien’) researcher in Computer Science has brought unique challenges to my academic career. Through my own experience, I began to understand the larger set of issues in the education system caused by lack of representation, racial and gender stereotype threat, and xenophobia. I also believe that these challenging experience led me to be more empathetic towards others’ challenges.

- **Lack of certain opportunities:** I left my family and friends behind in India in pursuit of better opportunities; but I am ineligible for so many opportunities such as fellowships, training programs, grants, postdoctoral and research positions. This leaves me feeling that due to where I was born, my research and contributions are considered less valuable than others’ by the institutions underlying academic research.

  *Efforts to create opportunities open to all:* When I was excluded from certain programs specifically aiming to empower under-represented minorities in CS[^3], I emailed the program organizers urging them to work to change the eligibility criteria to include non-citizens. I understand that at times, such decisions are beyond an organizer’s control. As such, I would do everything in my power as a faculty to design programs open to all students regardless of race, gender, or nationality. If government funding disallows this, I will seek funding from industries and private donors to better serve under-represented minorities while petitioning for change at the university or governmental level.

- **Imposter Syndrome:** Imposter syndrome is known to affect women more adversely than men in STEM, oftentimes due to stereotype threat and lack of representation in higher levels of academia[^3]. At a top database conference where I presented my work, a senior and well-known researcher told me I should apply to their workplace because they want more women there, with no mention of my research. Experiences like this intensified my imposter syndrome, leaving me wondering how the interaction would have been different if I was a male.

  *Efforts to alleviate imposter syndrome:* What I found most helpful in dealing with imposter syndrome was attending seminars on how to tackle imposter syndrome, where I learned tangible exercises such as a “FeelGood” email folder composed of all the positive emails I received, and being part of a small cohort (3 female students) with senior-junior graduate students where we regularly discussed academic and non-academic challenges. As a faculty, I will organize such seminars and push for graduate cohorts to help students combat imposter syndrome.

- **Feeling unwelcome:** During the pandemic, while I was already struggling emotionally because my parents in India contracted Covid-19 and I could not travel to care for them, I and other international students were threatened by xenophobic and racist government policies[^2] that could have unexpectedly disallowed my stay in the US - a country to which I contribute both financially and intellectually. In the months the policy[^2] was being considered, I felt marginalized and unwelcome, increasing my stress and feelings of despair.

  *Efforts to support marginalized students:* The things that kept me going during such trying times were the parts of the academic community championing for all of us international students and having a supportive partner, friends, and advisors. While I understand that as a faculty I would be unable to directly control such policies, I can foster an inclusive environment for my lab, department, and university. My experiences will help me empathise with marginalized students and help be a support system for them when xenophobic policies are inevitably put forth in the future until our society at large can become more inclusive.

Having campaigned for Diversity, Equity, and Inclusion (DEI) and undertaken successful DEI efforts in my academic career so far, I will continue such endeavors in the future. As an Assistant Professor, I am committed to join or lead DEI programs and to strive for equitable opportunity both within and beyond the field of Computer Science. At a departmental level, I intend to organize research backed[^5] outreach programs to open up opportunities for underrepresented and underprivileged high school students and encourage them to pursue STEM fields. Successful outreach activities[^5] can include giving university tours to motivate students to apply for college or providing a platform where high school students interact with recent graduates who can vouch for the benefits of education. To obtain financial support to execute these outreach activities, I intend to apply for funds such as NSF CSforAll. Through such activities, I hope to instill the significance of education in breaking the cycles of whatever challenges they are facing. As a faculty member, I will recruit students not just on their objective abilities - such as whether they have published papers or their coding experience - but their abilities subjected to the challenges they faced and opportunities they had. Apart from considering a student’s individual history, I will also account for a student’s diversity in thought, culture, socioeconomic and educational background, and inter-disciplinary work. Moving forward, I envision a more diverse and more inclusive Computer Science and I am committed, heart and soul, to partake in individual and collective efforts to bring the change I envision.

[^3]: [https://faculty.umd.edu/presidentspostdoc/](https://faculty.umd.edu/presidentspostdoc/) or [https://wiscprof.engr.wisc.edu/](https://wiscprof.engr.wisc.edu/)