UC SANTA BARBARA

THE Current

March 23, 2021
Andrew Masuda

A Matter of Privacy

While conducting the 2020 census, the U.S. Census Bureau took a new approach to ensure that data collected from individuals and households remained confidential. They implemented a powerful new disclosure avoidance tool based on differential privacy (DP) to withstand modern privacy threats and protect data.

DP is a mathematical definition of privacy that provides provable guarantees against identifications of individual subjects in a dataset, while still allowing the dataset, as a whole, to be useful. In other words, anything an algorithm might output from a dataset containing an individual's information is just as likely to have come from a database that does not include said information. DP has proven valuable in scientific research, data-driven decision making, and the training of machine-learning models.

"Differential privacy has been widely adopted and become the de facto standard of privacy definition," said Yu-Xiang Wang, an assistant professor of computer science at UC Santa Barbara. "If you use an iPhone or Chrome browser, Apple and Google might be optimizing your experience using DP already to learn about the user community without learning about individuals in the community."

However, said Wang, the Eugene Aas Chair in Computer Science, DP is challenging to successfully implement in any practical setting because the algorithms often need to be tailor-made and privacy guarantees must be proven mathematically.

"The design and analysis of DP mechanisms are delicate and error-prone tasks even for experts," said Wang, who joined UCSB's faculty in July 2018. "There is no

standard in how the ‘privacy loss’ parameter should be chosen, interpreted and reported. The issue is amplified by the fact that the numerical value of the privacy loss parameter might not be an accurate measure.”

Wang has devised an innovative new approach to advance DP theory and develop algorithms, and ground his findings in concrete applications. His technique automates some of the complex mathematical derivations with numerical algorithms and computations, as well as provides new algorithms that publish comprehensive and private data-dependent reports. Now, his proposed project, “Exact Optimal and Data-Adaptive Algorithms and Tools for Differential Privacy,” has garnered support from the National Science Foundation (NSF) in the form of a prestigious Early CAREER Award, which comes with $500,000 over five years to support his research.

“It’s a great honor to receive a CAREER Award,” said Wang, who received his Ph.D. in statistics and machine learning form Carnegie Mellon University. “It is a recognition for the amazing work that my graduate students, in particular Yuqing Zhu, Rachel Redberg, and Chong Liu, have been doing on differential privacy over the last few years. The award also provides significant support to my lab and enables us to continue our effort to make differential privacy more accessible and useful in practice.”

“I offer sincere congratulations to Professor Wang on receiving this highly esteemed award,” said Rod Alferness, dean of the College of Engineering. “The recognition reflects the tremendous potential of his innovative research to create new approaches that address the privacy challenges and threats that must be overcome to unlock the full potential of artificial intelligence and big data technologies.”

Receiving a CAREER Award has special significance for Wang, who says the roots of his project trace back to his work at CMU with his late graduate advisor Stephen E. Fienberg, who died of cancer in 2016. His advisor’s wife, Joyce Fienberg, died tragically in 2018 during the Tree of Life mass shooting in Pittsburgh.

“This award means a lot to me personally, because I will be able to scale up the research as part of Dr. Fienberg’s legacy, knowing that Steve and Joyce would be thrilled to see the better world that this research brings about,” said Wang.

To address the existing challenges of DP, Wang plans to use a wide range of techniques from computing and probability to optimization theory and numerical analysis. To be pragmatic, he said, researchers need to find the best tools to solve a

problem and learn to use new tools whenever necessary.

“The combination of computing and math is especially powerful because computing will allow us to practically implement mathematical equations that are precise but not simple, rather than resorting to simple approximations,” said Wang. “DP is at a pivotal moment, transitioning from a theoretical construct into a practical technology. Our research eases the growing pains and paves the way for DP to be used and deployed in a wider array of applications.”

Wang’s research has far-reaching implications beyond helping the U.S. Census provide more accurate and private data and model releases. His group is developing applications through a number of collaborative efforts, including one with researchers at UC San Diego to apply DP to clinical research studies. They also are working with Evidation Health, an industry leader that collects and analyzes behavior and healthcare information through wearable devices.

“Evidation Health has been supporting our DP research through their partnership with the university’s Center for Responsible Machine Learning,” said Wang, who is co-director of the center. “Our goal is to develop efficient DP algorithms for person-generated health data that will enable data sharing in a way that preserves patient privacy and opens doors for researchers who want to access valuable data for the greater good.”

Wang’s proposal includes several educational components as well. He will use the findings to expand his open-source software library, autodp, which stands for Automating Differential Privacy Computation. Autodp provides an online resource to make state-of-the-art differentially private computations more accessible and provide anyone with a hands-on learning experience in DP.

He also plans to train and mentor the next generation of computer scientists through the Computer Science Department’s Early Research Scholars Program (ERSP). Funded by the NSF, the year-long research apprenticeship program targets women and underrepresented minority students in their second year of study, providing them with research experience. Wang has proposed multiple research projects that involve undergraduate students applying DP theory into practice.

Tags
Artificial Intelligence

About UC Santa Barbara

The University of California, Santa Barbara is a leading research institution that also provides a comprehensive liberal arts learning experience. Our academic community of faculty, students, and staff is characterized by a culture of interdisciplinary collaboration that is responsive to the needs of our multicultural and global society. All of this takes place within a living and learning environment like no other, as we draw inspiration from the beauty and resources of our extraordinary location at the edge of the Pacific Ocean.