Georgia Tech Research Corporation

Technologies
Available for LICENSING

OFFICE OF TECHNOLOGY LICENSING

https://licensing.research.gatech.edu | techlicensing@gtrc.gatech.edu

Enabling Privacy-Preserving Search Over Fuzzy Databases

New protocols to improve biometric security while maintaining efficiency

This new system can be used to improve the privacy of facial recognition searches and other forms of biometric-based surveillance and identification (e.g., voice, iris, fingerprint etc.) while maintaining efficiency for demanding applications. Georgia Tech's innovation enables users to query large databases that contain biometric data from real-time surveillance in such a way that only the identities of those captured by surveillance sensors and are in the database will be revealed. The system keeps private the identities of everyone else in the surveillance data. In addition, even when a user learns the identity of a person in the surveillance data, the system specifies that the server will learn nothing about either the query or the result thanks to its quantum-safe cryptographic design.

Georgia Tech makes this capability possible with the introduction of two new protocols—fuzzy labeled set intersection (FLPSI) and its extension, batch-FLPSI (BFLPSI). FLPSI addresses the gap in current privacy-preserving database search technologies that do not accommodate search over fuzzy data such as biometrics. It efficiently computes the intersection of noisy input sets by considering closeness/similarity rather than exact matches. It is the first protocol of its kind to achieve sublinear communication cost relative to a database, thereby achieving efficiency important for databases containing very large numbers of records. Efficiency is further improved with the use of batch-FLPSI, which makes multiple FLPSI queries at the same time at a cost similar to that of a single query.

Summary Bullets

- **Privacy-preserving**: Improves the security of biometric-based surveillance, identification, or searches for individuals using their biometric data over private databases
- **Timely**: Addresses the requirements of recent and emerging privacy protection regulations and policies
- **Practical**: Fills a gap in privacy-preserving search technologies, which currently do not accommodate searching of fuzzy data such as biometrics

Solution Advantages

- **Privacy-preserving**: Improves the security of biometric-based surveillance, identification, or searches for individuals using their biometric data over private databases
- **Timely**: Addresses the requirements of recent and emerging privacy protection regulations and policies

- **Practical**: Fills a gap in privacy-preserving search technologies, which currently do not accommodate searching of fuzzy data such as biometrics
- **Efficient**: Offers the first protocol of its kind to achieve sublinear communication cost relative to a database, and outperforms querying speeds compared to alternative protocols without the need for a high-speed network connection
- **Economical**: Decreases communication costs by up to 611x by using FLPSI and improves querying speeds up to 129x compared to the state-of-the-art among alternative protocols with further querying speed improvements of up to 13x with the use of BFLPSI
- **Compatible**: Is installed as a software plugin compatible with existing surveillance infrastructure, fuzzy databases, and network connections—effectively sidestepping the need for additional (or sophisticated) computation power or sensors

Potential Commercial Applications

Georgia Tech’s technology is applicable to biometric search, identification, and surveillance, especially those requiring improved privacy and high efficiency, including among others:

- “Person of interest” searches by law enforcement
- Detection of fake or “shadow” social media accounts
- Shoplifter identification and biometric payment in retail settings
- Background checks
- Privacy-preserving medical records searches

Private businesses and law enforcement agencies that need to comply with regulations and policies prohibiting biometric surveillance due to security concerns may be able to resume such surveillance while remaining in compliance with regulations by employing this technology.

Background and More Information

Recent advances in deep learning-based biometric identification have made possible real-time identification of individuals in surveillance footage. While potentially beneficial to public safety, indiscriminatory identification of people in these videos and surveillance data raises serious privacy concerns and has become subject to regulations. While other privacy-preserving database search protocols are available, they do not offer search capabilities for fuzzy data, which includes biometrics. Therefore, in preserving privacy, these solutions make it difficult to use biometric data in a useful way. Other approaches also focus on exact matches and are inefficient for large databases. Georgia Tech’s innovation addresses these shortcomings with a solution that makes use of biometric data in a specific and practical way while preserving privacy and maintaining cost effective and efficient operation even for databases containing tens of millions of records.

Inventors

- Erkam Uzun
  PhD Student - Georgia Tech Computer Science
- Pak Chung
  Research Scientist - Georgia Institute of Technology
- Dr. Vladimir Kolesnikov
  Associate Professor - Georgia Tech School of Computer Science
- Alexandra Boldyreva
  Associate Professor – Georgia Tech College of Computing

- Dr. Wenke Lee
  John P. Imlay, Jr., Chair of Software - Professor - Georgia Tech College of Computing

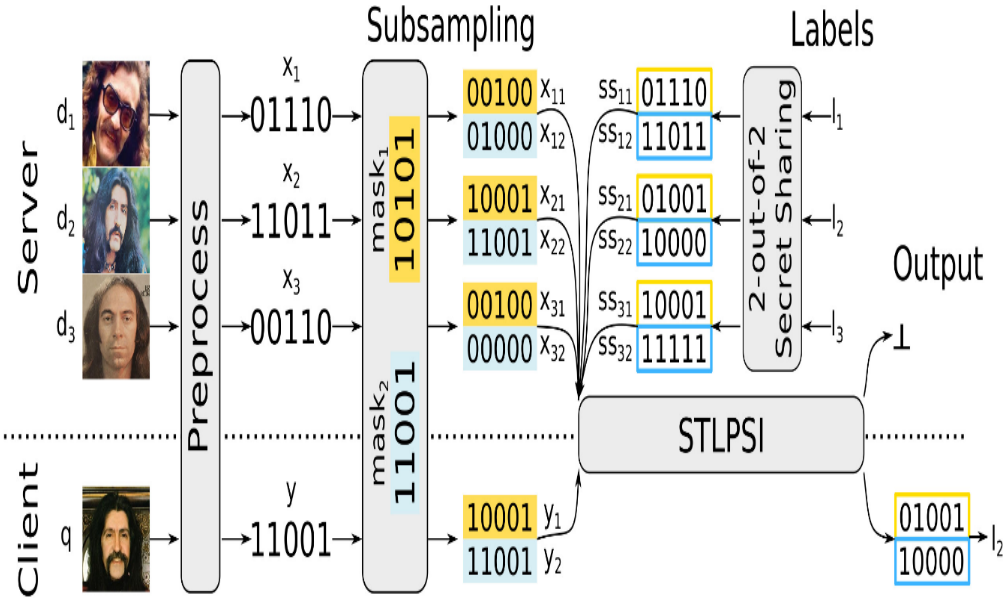
IP Status

: 63/052,201

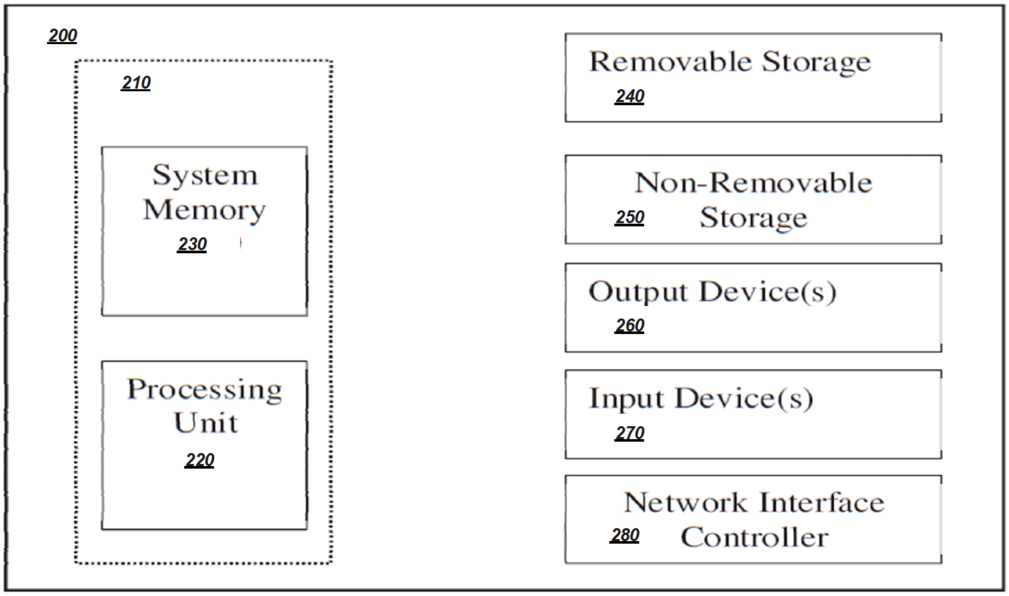
Publications

Facial Recognition: How to address privacy concerns, YouTube - March 16, 2021

Images



Prior to performing a query, this technology is configured to perform pre-processing to transform raw biometric inputs (e.g., facial photos, video, voice, DNA information, etc.) into bit vectors that can be readily evaluated for “closeness.” Then, a sub-protocol called "Set Threshold LPSI (STLPSI)" is used for searching exactly matching subsamples of these bit vectors to secretly transfer a matching record's label to the client.



This computer architecture is an example of a computer system capable of executing the software components for Georgia Tech’s FLPSI and/or batch-FLPSI protocols.

Visit the Technology here:
Enabling Privacy-Preserving Search Over Fuzzy Databases