Research Statement - Brian Jay Tang

Motivation

Privacy, while fundamental to freedom of expression and freedom from oppression, has become a fleeting value. With emerging artificial intelligence (AI) technologies, privacy is now left with an uncertain future. At the heart of privacy is protecting users from harm and informing them about the collection and processing of their data. Users should know if and when a service/website misuses their data and be able to protect themselves. Legal measures like the GDPR and CCPA have been instituted for privacy protection and AI regulation, yet regulators face the task of keeping pace with rapid technological advancements. With AI evolving rapidly, service providers bear the brunt of the responsibility to self-regulate and uphold user privacy. The verification of privacy practices' correctness, consistency, and legality requires extensive debugging, network traffic analysis, and consultation with privacy experts.

Therefore, my objectives are to (1) explore unique ethical and privacy concerns resulting from potential AI misuse, and (2) develop accessible tools to empower stakeholders in safeguarding user data privacy.

Past Projects

In previous work, I sought to address the issues of correctness and transparency in privacy policies, detecting violations in areas such as cookie settings [1], browser extensions [2], and opt-out mechanisms [3]. We found many platforms infringing upon statutes within the GDPR, UK-DPA, FTC Act, and CCPA. For instance, numerous Chrome extensions breached Chrome's developer policies, collecting user data without proper disclosure. Additionally, we found that third-party tracker opt-out mechanisms often failed to stop data collection activities. I have also contributed to the creation of user-centric privacy tools, such as an anti-face recognition system [4], and I created a patented software-based privacy film designed to protect smartphone information from prying eyes [5]. Furthermore, I developed a framework for social robots to maintain user privacy during interactions [6]. Lastly, my research has encompassed the examination of safety, security, and fairness issues machine learning systems [7, 8].

Ongoing and Future Projects

The advent of large language models (LLMs) has complicated privacy by blurring the boundaries between data usage purposes. These large transformer models are general-use data processors, enabling them to interpret images, make sensitive inferences about users from various data sources, and make decisions as embodied agents. A user agreeing to data collection for the purposes of "marketing and online behavioral advertisement", should not include using LLMs to infer users' psychological/emotional states or users' sensitive traits for the purpose of advertising.

We are currently examining risks associated with the use of large language models for advertising and data processing, focusing on user perceptions of personalized advertisements and user profiles generated by these models [9]. Additionally, we are investigating augmenting user privacy controls with AI, using an automated framework for modeling website contexts, executing UI actions, and monitoring context-specific data collection [10].

My future projects focus on LLM systems' integration into cyber-physical systems such as autonomous vehicles, smart glasses, and IoT-integrated buildings. For example, buying cars or smart glasses should not implicitly give companies the right to feed one's camera data into GPT-4's vision transformer to infer a user's behaviors or interests. Irrespective of these concerns, AI systems continue to develop rapidly, outpacing privacy regulations and privacy enhancing technologies. I aim to prevent this type of misuse by creating protection and opt-out mechanisms that scale with cyber-physical systems.

Brian Jay Tang
(630) 880 3691 • bjaytang@umich.edu

References

[1] **Brian Tang**, D. Bui, and K. G. Shin, "Detection and analysis of cookie violations," in *In Review: 30th ACM Conference on Computer and Communications Security*, 2023.

[2] D. Bui, **Brian Tang**, and K. G. Shin, "Detection of inconsistencies in privacy practices of browser extensions," in *44th IEEE Symposium on Security and Privacy*, 2023.

[3] ——, "Do opt-outs really opt me out," in *29th ACM Conference on Computer and Communications Security 2022*, 2022.

[4] V. Chandrasekaran, C. Gao, **Brian Tang**, K. Fawaz, S. Jha, and S. Banerjee, "Face-off: Adversarial face obfuscation," in *21st Privacy Enhancing Technologies Symposium*, 2021. [Online]. Available: `https://arxiv.org/abs/2003.08861`.

[5] **Brian Tang** and K. G. Shin, "Real-time protection of mobile device screen information from shoulder surfing," in *Major Revision: 32nd USENIX Security Symposium*, 2023.

[6] **Brian Tang**, D. Sullivan, B. Cagiltay, V. Chandrasekaran, K. Fawaz, and B. Mutlu, "Confidant: A privacy controller for social robots," in *17th ACM/IEEE International Conference on Human-Robot Interaction*, 2022. [Online]. Available: `https://arxiv.org/abs/2201.02712`.

[7] V. Chandrasekaran, **Brian Tang**, N. Papernot, K. Fawaz, S. Jha, and X. Wu, "Rearchitecting classification frameworks for increased robustness," 2020. arXiv: `1905.10900`. [Online]. Available: `https://arxiv.org/abs/1905.10900`.

[8] H. Rosenberg, **Brian Tang**, K. Fawaz, and S. Jha, "Fairness properties of face recognition and obfuscation systems," in *32nd USENIX Security Symposium*, 2023. [Online]. Available: `https://arxiv.org/abs/2108.02707`.

[9] **Brian Tang**, N. T. Curran, F. Schaub, and K. G. Shin, "Embedding advertising in llm chatbots: Risks and ethical considerations," in *In Preparation: ACM CHI Conference on Human Factors in Computing Systems*, 2025.

[10] **Brian Tang** and K. G. Shin, "Steward: Natural language web automation," in *Under Submission: The 30th Symposium on Operating Systems Principles*, 2024.