The paper represents a critical result in our understanding of the privacy risks embedded in privacy-enhancing technologies (PETs) themselves—specifically, second-generation Compromised Credential Checking (C3) systems such as “Might I Get Pwned” (MIGP). The previous generation of compromised credential checking services, e.g., Have I Been Pwned (HIBP), revolutionized the field by centralizing massive breach collections and offering accessible, real-time breach exposure checks to both users and enterprises.

While MIGP was designed to defend against credential tweaking attacks in a privacy-preserving manner, the authors identify and formalize subtle, yet devastating, cryptographic leakage within the protocol. This leakage enables breach extraction attacks, whereby an honest-but-curious client can reconstruct sensitive and potentially non-public breached credentials stored on the server—directly undermining the very privacy MIGP was intended to protect.

This work's methodology is very novel. It provides the first formal leakage profile of MIGP, introduces a taxonomy of novel leakage phenomena such as τ-collisions, and models how these interactions can be exploited through carefully constructed attacks. This paper goes beyond theoretical insight: it builds and evaluates several practical attacks, including one that requires knowledge of just a single password and another that works with no prior knowledge of the target’s credentials. Notably, its graph-based neural network models infer structural “templates” of unknown passwords purely from observable leakage—a creative and powerful demonstration of how adversaries can abuse even minimal leakage.

In addition to the depth of its cryptanalysis, the team proactively informed Cloudflare of the vulnerabilities, whose acknowledgment and assessment underscore the real-world relevance and urgency of their findings. It also proposes MIGP 2.0, a revised protocol that eliminates the identified leakage while retaining the core functionality of privacy-preserving credential checking.

Finally, this work was one of the finalists for the "Best Cryptographic Attack" category at Pwnie Awards 2024.