Cryptographic watermarking for preventing deepfakes in media data

Graduate Internship (Master 2), Summer 2025

Location	PQShield SAS, Paris 8 Rue des Pirogues de Bercy (WeWork Bercy)
Contact	Thomas Prest, Thomas Espitau, Pierre-Yves Strub, Adrian Thillard
Starting date and duration	Q2 2025 (flexible), 6 months
Gross salary	1350 € / month

Internship

Background

The proliferation of generative AI technologies has significantly compromised the authenticity of digital images, particularly through the creation of *deepfakes*. This internship aims to develop and enhance cryptographic watermarking techniques to verify image authenticity and provenance, drawing on recent advancements in zero-knowledge proofs (ZKPs) and image manipulation verification

Scope of the internship

The primary objective of this internship is to explore and implement watermarking methodologies that can withstand deepfake manipulations. Key goals include:

1. **Understand verification of Image Transformations**: Utilize zkSNARKs to prove the authenticity of image transformations, as outlined in the recent **VIMz** [1] and **VerITAS** [2], **ZK-IMG** [4] frameworks.
2. **Watermarking Framework Development**: Design a watermarking system that maintains the integrity of original images while allowing for the verification of authorized transformations. We need to ensure implement privacy-preserving proof systems to guarantee that *only allowed* transformations are applied.

To do so, the intern will undertake the following tasks:

- **Literature Review**: Conduct a comprehensive analysis of current watermarking and ZKP approaches, focusing on challenges in proving image authenticity.
- **Framework Development**: Design and prototype watermarking algorithms, for instance inspired by the low prover complexity methods demonstrated in **VIMz** [1].
- **Performance Evaluation**: Test the watermarking systems against increasingly large images and various transformations. This implies proposing a proof-of-concept implementation and designing a test suite for it.

About PQShield

PQShield is a cybersecurity scaleup that specialises in post-quantum cryptography, protecting information from today's attacks while readying organisations for the threat landscape of tomorrow. It demonstrates quantum-safe cryptography on chips, in applications and in the cloud. We are headquartered in Oxford, with additional teams in the Netherlands, Germany and France.

PQShield SAS, based in Paris (France), concentrates the research activities of PQShield. Our mission is to come up with innovative algorithmic and/or protocol-level solutions to real-world cryptographic problems. Besides post-quantum cryptographic primitives, our research interests include implementation security and advanced cryptosystems and protocols such as secure messaging, threshold schemes, and multiparty computation.

Requirements and how to apply

Ideal candidates should possess:

- A solid understanding of cryptographic principles, particularly in ZKPs and digital signatures.
- Proficiency in programming languages relevant to cryptographic algorithm implementation (at least Python, better if C also)
- Strong collaboration skills and the ability to communicate clearly.

Supervision and Mentorship: Interns will be mentored by the researchers of the cryptography R&D team of PQShield in Paris, with regular progress meetings to facilitate knowledge exchange.

To apply, please send your CV to the contact points listed at the top of this document.

References

1. Dziembowski, S., Ebrahimi, S., & Hassanizadeh, P. (2023). VIMz: Verifiable Image Manipulation using Folding-based zkSNARKs.
2. Datta, T., Chen, B., & Boneh, D. (2023). VerITAS: Verifying Image Transformations at Scale.
3. Della Monica, P., Visconti, I., Vitaletti, A., & Zecchini, M. (2023). Trust Nobody: Privacy-Preserving Proofs for Edited Photos with Your Laptop.
4. Kang, D., Hashimoto, T., Stoica, I., & Sun, Y. (2023). ZK-IMG: Attested Images via Zero-Knowledge Proofs to Fight Disinformation.