Balancing Privacy and Accountability on Encrypted Messaging Platforms

Alistair Pattison and Nicholas Hopper

This work introduces a protocol in which **the sender of an abusive message can be verified if there is sufficient agreement among a group of moderators**. The protocol **retains all security properties of the messaging service for unreported messages**.

The Protocol

This work extends a protocol by Issa, Alhaddad, and Varia called Hecate. In the Hecate protocol, every message sent via an encrypted service requires a token containing (among other things), the encryption

$$x_1 \coloneqq \textsc{Enc}_{mod}(id_{src}).$$

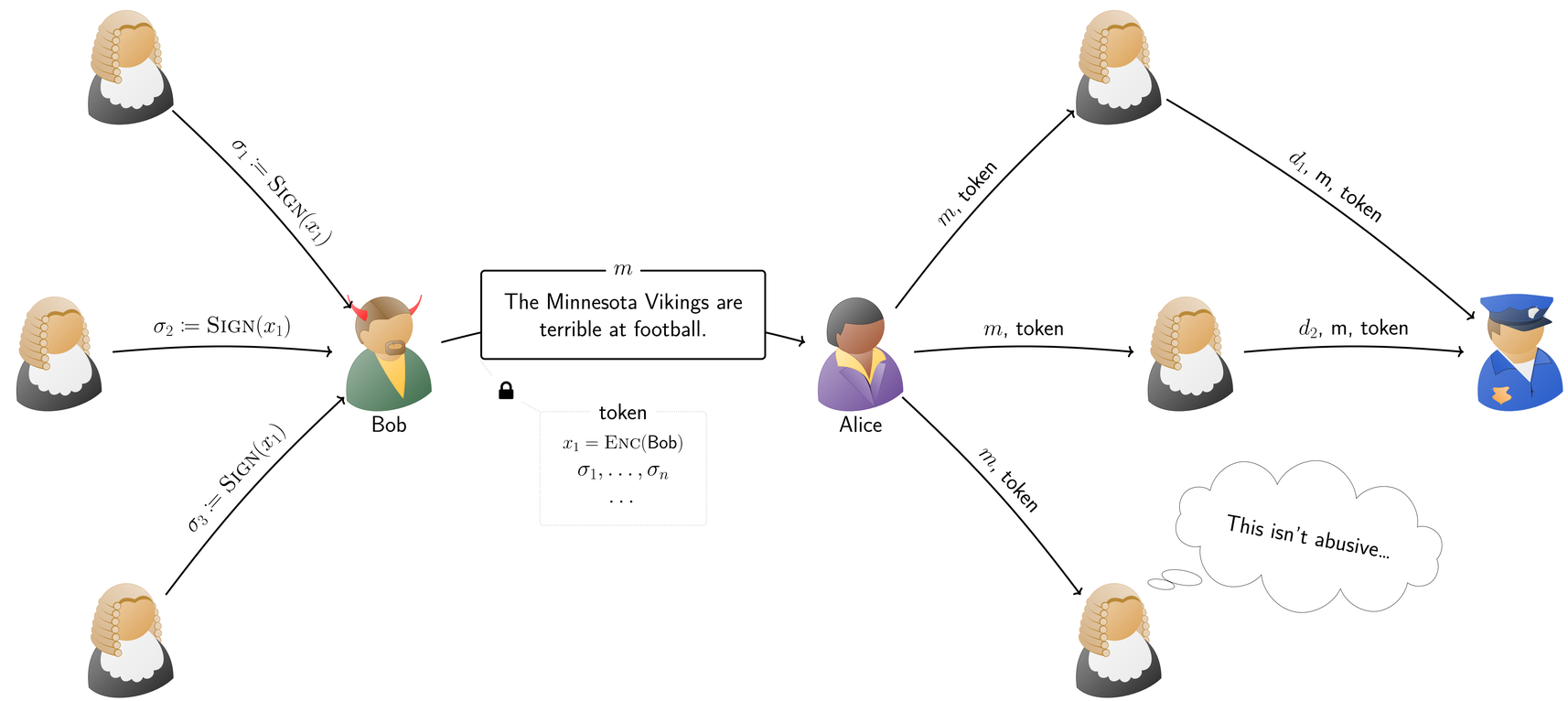
This x_1 is attached to a sent message so that if the message is reported, a moderator can decrypt x_1 to recover the sender's identity. Using threshold cryptography, this work extends Hecate so that the source's identity is revealed only under sufficient agreement from a group of moderators.

1: **function** $\textsc{CreateToken}(id_{usr})$
2: Compute $x_1 = \textsc{Enc}(id_{src})$ using a k out of n threshold encryption scheme.
3: Package x_1 into a token and send a signature request to each moderator.
4: Each moderator verifies verifies x_1 and returns their signature share on the token.
5: Once sufficient responses are received, combine the shares into into a valid signature.

1: **function** $\textsc{HandleReport}(m,\ x_1)$
2: A user sends a request to all n moderators containing the reported message m and the encrypted id, $x_1 = \langle c_1, c_2 \rangle$.
3: If a moderator believes that the message is worth acting upon, she forwards the decryption share d_i to the appropriate authorities.
4: If enough decryption shares are received, one combines the decryption shares to recover x_1.

Implementation

We implement the protocol using Rust and run each party in a separate container communicating over HTTP. We estimate the cost of adoption to be around around $100 per day for the entirety of WhatsApp. See the poster abstract for benchmark results.



Poster Abstract **Implementation** **Hecate paper**