



MP3: A BETTER PRIVACY-PRESERVING PRESENCE PROTOCOL

{ RAHUL PARHI, MICHAEL SCHLIEP, NICHOLAS HOPPER } UNIVERSITY OF MINNESOTA



INTRODUCTION

Online communication is very prevalent in the day to day lives of nearly everyone. To “connect” with one’s friends, knowing when they are online is extremely important. All services with these sort of “buddy” lists know the entire graph structure of the social network. This is an invasion of privacy. What if who you are friends with is sensitive information?

What’s the solution?

Using the magic of cryptography, this problem can be solved. In mid-2015, Borisov, Danezis, and Goldberg proposed DP5—the Dagstuhl Privacy Preserving Presence Protocol P—a privacy-preserving cryptographic protocol [1]. Though DP5 offers a solution to this problem, there are optimizations and improvements to be made. We propose MP3—the Minneapolis Private Presence Protocol—as an improvement to DP5.

MP3 OVERVIEW

At a higher level, MP3 works as follows. Alice will upload her encrypted presence to the registration server, and Bob will request Alice’s presence from the lookup server. Lookup is done using Private Information Retrieval (PIR) [2]. That is, the server will not know that Bob requested Alice’s presence. Thus, no information of the social graph is leaked. To use PIR, time is divided into epochs, long (T_J) and short (t_j). For registration and lookup, MP3 makes use of a dynamic broadcast encryption scheme and bilinear maps [3] for encrypting and decrypting entries sent to the server.

BILINEAR MAPS

Let G_1 and G_2 be two additive (written multiplicatively) cyclic groups of prime order p and G_T be a multiplicative group of order p . Define a map $e : G_1 \times G_2 \rightarrow G_T$ where, $\forall g_1 \in G_1, g_2 \in G_2$, and $a, b \in \mathbb{Z}/p\mathbb{Z}$ that:

$$e(g_1^a, g_2^b) = e(g_1, g_2)^{ab} \in G_T \quad (1)$$

Call this function a pairing function. This is vital for encryption and decryption.

REFERENCES

- [1] Nikita Borisov, George Danezis, and Ian Goldberg. DP5: A private presence service. *Proceedings on Privacy Enhancing Technologies*, 2015(2):4–24, 2015.
- [2] Casey Devet and Ian Goldberg. The best of both worlds: Combining information-theoretic and computational pir for communication efficiency. In *Privacy Enhancing Technologies*, pages 63–82. Springer, 2014.
- [3] Cécile Delerablée, Pascal Paillier, and David Pointcheval. Fully collusion secure dynamic broadcast encryption with constant-size ciphertexts or decryption keys. In *Pairing-Based Cryptography—Pairing 2007*, pages 39–59. Springer, 2007.

SETUP & LONG-TERM EPOCH

To participate in MP3, Alice randomly selects $G \in G_1$, $H \in G_2$, and $\gamma \in \mathbb{Z}/p\mathbb{Z}$ and stores (G, H, γ) as her *manager key*. For every friend i , she shares a *decryption key* with them out-of-band as:

$$x_i \in \mathbb{Z}/p\mathbb{Z}, \quad A_i = G^{\frac{x_i}{\gamma+x_i}}, \quad B_i = H^{\frac{1}{\gamma+x_i}}$$

and her current long-term presence key, P_a^J .

Long-term Registration

During epoch T_{J-1} , Alice computes:
Generates new P_a^J (from Ed25519)

$$\begin{aligned} x_j \in \mathbb{Z}/p\mathbb{Z} & \quad P_a^j = g_1^x \\ x_r^J \in \mathbb{Z}/p\mathbb{Z} & \quad B_r^J = H^{\frac{1}{\gamma+x_r}} \\ \text{Update } H = B_r^J & \quad k^J \in \mathbb{Z}/p\mathbb{Z} \\ C_1^J = G^{k\gamma} & \quad C_2^J = H^{k\gamma} \\ K^J = e(G, H)^k & \quad C_a^J = \text{AEAD}_K^0(P_a^J || P_a^j) \end{aligned}$$

$$\sigma^J = \text{sign}(P_a^{J-1}, x_r^J, B_r^J, C_1^J, C_2^J, C_a^J)$$

$$\text{Upload } P_a^{J-1}, x_r^J, B_r^J, C_1^J, C_2^J, C_a^J, \sigma^J$$

Long-term Lookup

During epoch T_J , Bob requests from the PIR servers the entry beginning with P_a^{J-1} , then updates his B to $\left(\frac{B_r^J}{B}\right)^{\frac{1}{x-x_r}}$, computes $K^J = e(G, H)^{k^J} = e(C_1^J, B) \cdot e(A, C_2^J)$, and use K^J to decrypt C_a^J , retrieving P_a^J and P_a^j .

IMPROVEMENTS TO DP5

The main bottleneck in DP5 is the scaling with large user bases. This comes from the long-term epoch being very expensive. MP3 solves just that. The main contribution of MP3 is a smaller long-term database, and, as a side-effect, less CPU operations and less bandwidth required to run. This effectively reduces the cost of running this privacy-preserving presence mechanism. There is a loss in privacy, though, in the fact that whenever you revoke a friend, they know you revoked them, while in DP5, a revoked friend just sees it as you are never online. MP3 also allows for (essentially) infinite friends, whereas DP5 had a max number of friends.

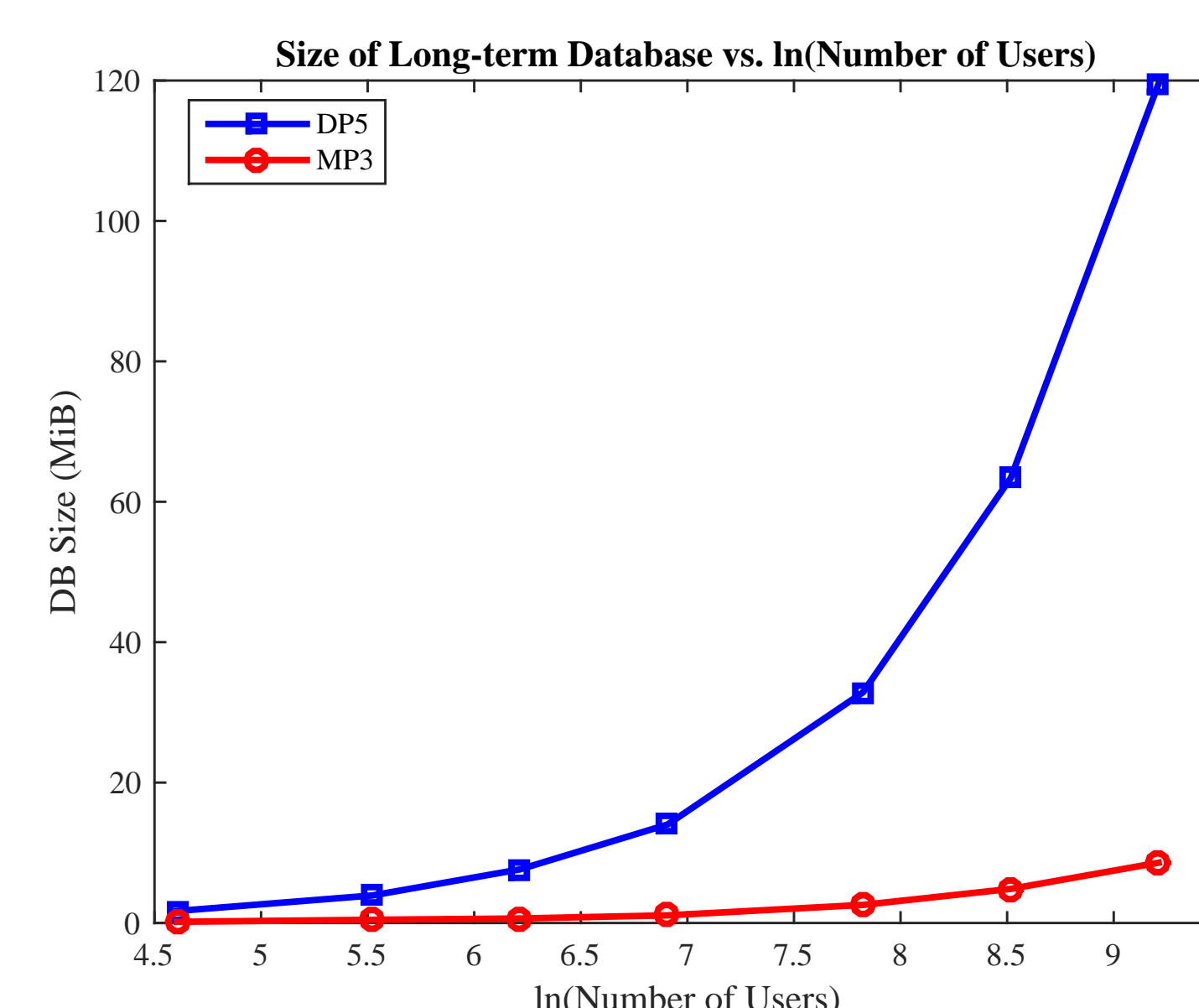


Figure 1: Semi-log plot of how the long-term database scales with the number of users for DP5 and MP3. Number of users ranges from 100 to 10,000.

Table 1: Sizes of long-term databases. $N = 1000$ (for both) and $N_{fmax} = 100$ (for DP5).

DP5		
	Req	Resp
DB Size	14.04 MiB	
Registration	9004 B	5 B
MP3		
	Req	Resp
DB Size	1.09 MiB	
Registration	324 B	5 B

SHORT-TERM EPOCH

Short-term Registration

During epoch t_{j-1} , Alice computes:
 $s_a^j = H_1(t_j)^{x_j}$ $K_a^j = \text{PRF}_{H_3(P_a^j)}(t_j)$
 $m_a^j = \langle \text{presence msg} \rangle_j$ $c_a^j = \text{AEAD}_{K_a^j}^j(m_a^j)$
sends (s_a^j, c_a^j) to server, server computes:
 $\text{ID}_a^j = H_0(e(g_1, s_a^j)) = H_0(e(P_a^j, H_1(t_j)))$
and sends (ID_a^j, c_a^j) to short-term user database and (ID_a^j, s_a^j) to short-term signature database.

Short-term Lookup

Bob computes $\text{ID}_a^j = H_0(e(P_a^j, H_1(t_j)))$ and requests the entry associated with that ID from the PIR servers. He receives back c_a^j . Bob can compute $K_a^j = \text{PRF}_{H_3(P_a^j)}(t_j)$. Bob can finally use K_a^j to decrypt c_a^j and retrieve m_a^j , Alice’s current online presence.

PRIVACY

Pros:

- Only your friends know your online status
- Perfect forward secrecy is maintained in the event of a compromise
- Social graph is completely unknown
- Your identity is completely anonymous

Cons:

- Revocations are explicit
- Can no longer temporarily suspend a friend

CONCLUSION

We propose MP3, and efficient privacy-preserving presence protocol that allows for cheaper operation than its counterpart, DP5. The long-term database was reduced in size and computational time by using a dynamic broadcast encryption scheme, thus decreasing total cost of operation. This optimization comes at the cost of a loss in privacy in that revocations are explicit and temporary suspensions are no longer possible.