Provably-secure symmetric private information retrieval with quantum cryptography

Wen Yu Kon[1], Charles Ci Wen Lim[1,2]

[1]Department of Electrical & Computer Engineering, National University of Singapore

[2]Centre for Quantum Technologies, National University of Singapore

Introduction

Private information retrieval (PIR) is a database query technique which guarantees user privacy, where the user can learn a particular entry of the database of his interest, but his query would be hidden from the data centre[1]. Symmetric private information retrieval (SPIR) takes PIR further by additionally offering database privacy, where the user cannot learn any additional entries of the database[2]. Unconditionally secure SPIR solutions with multiple databases are known classically but are unrealistic because they require long shared secret keys between the parties for secure communication and shared randomness. Here, we propose using quantum key distribution[3] (QKD) instead for a practical implementation, which can realise both the secure communication and shared randomness requirements.

Classical SPIR Protocol

A multi-database SPIR protocol has a user U, who interacts with k data centres D_j each having a copy of the database $w = (w_1, \dots, w_n)$. The user desires to learn entry x and has a source of local randomness R. The protocol starts with the user querying the database using query functions f_{query}, followed by replies from the database using answer functions f_{ans}, which will be used by the user to decode $\hat{w}_x$. At the end of the protocol, each parties' collection of bits is referred to as the view, V.

An SPIR protocol satisfies three condition[2] (informally):

1. Correctness: *User gets his desired entry,* $\hat{w}_x = w_x$.
2. User Privacy: *Each data centre's view* V_{D_j} *is independent of* x.
3. Database Privacy: *The user's view* V_U *is independent of* w, *except perhaps for some* $w_{x'}$.

One-Round SPIR Protocol				
Step	D_1		U	D_2
Input:	w		$R,\ x$	w
Key pair ($D_1 \leftrightarrow D_2$):	K_5	$\longleftrightarrow$		K_6
Key pair ($U \leftrightarrow D_1$):	K_1	$\longleftrightarrow$	K_2	
Key pair ($U \leftrightarrow D_2$):			K_4	$\longleftrightarrow$ K_3
Query:			$Q_1 = f_{\text{query},1}(x,R),\quad Q_2 = f_{\text{query},2}(x,R)$	
OTP ($U \to D_1$):	$\tilde{Q}_1 = C_{Q_1} \oplus K_1^{\text{dec}}$	$\xleftarrow{\mathscr{C}_{U1}}$	$C_{Q_1} = Q_1 \oplus K_2^{\text{enc}}$	
OTP ($U \to D_2$):			$C_{Q_2} = Q_2 \oplus K_4^{\text{enc}}$	$\xrightarrow{\mathscr{C}_{U2}}$ $\tilde{Q}_2 = C_{Q_2} \oplus K_3^{\text{dec}}$
Answer:	$A_1 = f_{\text{ans},1}(\tilde{Q}_1, w, K_5)$			$A_2 = f_{\text{ans},2}(\tilde{Q}_2, w, K_6)$
OTP ($D_1 \to U$):	$C_{A_1} = A_1 \oplus K_1^{\text{enc}}$	$\xrightarrow{\mathscr{C}_{U1}}$	$\tilde{A}_1 = C_{A_1} \oplus K_2^{\text{dec}}$	
OTP ($D_2 \to U$):			$\tilde{A}_2 = C_{A_2} \oplus K_4^{\text{dec}}$	$\xleftarrow{\mathscr{C}_{U2}}$ $C_{A_2} = A_2 \oplus K_3^{\text{enc}}$
Decoding:			$\hat{w}_x = f_{\text{dec}}(\tilde{A}_1, \tilde{A}_2, Q_1, Q_2, x, R)$	

Table 1: Generic one-round two-database SPIR protocol

QKD SPIR Protocol

To analyse SPIR protocols that utilise QKD keys, it is necessary to generalise the original SPIR security definition. We say that a SPIR is $(\eta_{corr}, \eta_{UP}, \eta_{DP}, \eta_{PS})$-secure if it satisfies

1. Correctness: *Assuming the user and the data centres are honest, then for any* x *and* w, *the protocol must fulfil* $(1 - p_{fail})\Pr[\hat{w}_x \neq w_x | pass] \leq \eta_{corr}$.
2. User Privacy: *Assuming the user is honest, then for any* D_j *and* E, *their total view satisfies* $\Delta\left(\rho_{D_jE}(x), \rho_{D_jE}(x')\right) \leq \eta_{UP}$ for all x and x'.
3. Database Privacy: *Assuming the data centres are honest, then for any* U *and* E, *and for any* x, *there exist an* x' *such that for all* w *and* w' *with* $w_{x'} = w'_{x'}$, *their total view satisfies* $\Delta\left(\rho_{UE}(w), \rho_{UE}(w')\right) \leq \eta_{DP}$.
4. Protocol Secrecy: *Assuming the user and the data centres are honest, then for any* E, *her view satisfies* $\Delta\left(\rho_E(x,w), \rho_E(x',w')\right) \leq \eta_{PS}$ for all (x,w) and (x',w').

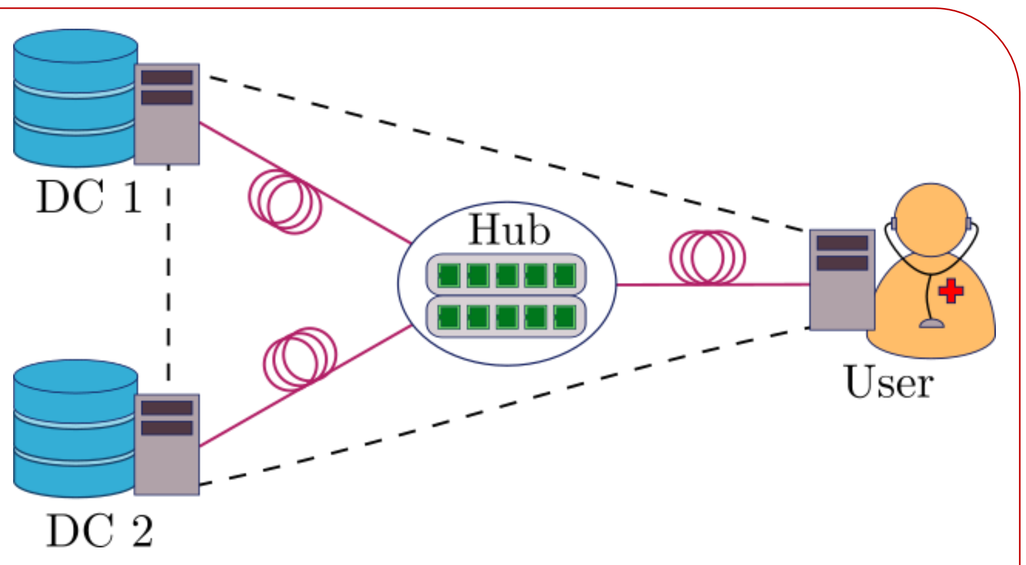

Figure 1: Schematic of a MDI-QKD network which can supply keys for the SPIR protocol. The central node (hub) connects to the user and two data centres with optical fibre (solid lines). Using the physical connection, any two parties in the protocol can establish a secure QKD link (dotted lines) via the central node

Using this modified definition, we prove that QKD keys can indeed be used to facilitate multi-database SPIR implementation, using schemes such as that shown in Fig. 1. This is summarised in the below (Note that classical SPIR protocols are $(0,0,0,0)$-secure.

Theorem 1. *A two-database* $(0,0,0,0)$*-secure SPIR protocol using* ε*-secure QKD keys in place of ideal keys, where* $\varepsilon = \varepsilon_{corr} + \varepsilon_{sec}$, *is* $(3\varepsilon_{corr}, 2\varepsilon, 2\varepsilon, 4\varepsilon)$*-secure.*

Numerical Simulation

We perform a numerical simulation for SPIR using MDI-QKD with decoy states focused on communication between the user and data centres[4]. For a database of n entries and L entry size, Fig. 2 shows the number of signals required, N, for SPIR, for different metropolitan distances. Included are four scenarios:

- iTunes: 60 million songs, about 10MB in size each
- Electronic Health Records: 5.7 million patients, about 5MB medical charts[5]
- Fingerprint Data: 7.7 billion world population, 500 bytes for each minutiae data[6]
- Genome: 19116 genes, with up to 9.88 million bits for both alleles[7]

We also included a protocol that considers a relaxed version of SPIR where the user is only allowed learn a single bit – which includes values like $w_x \oplus w_{x'}$(blue line).

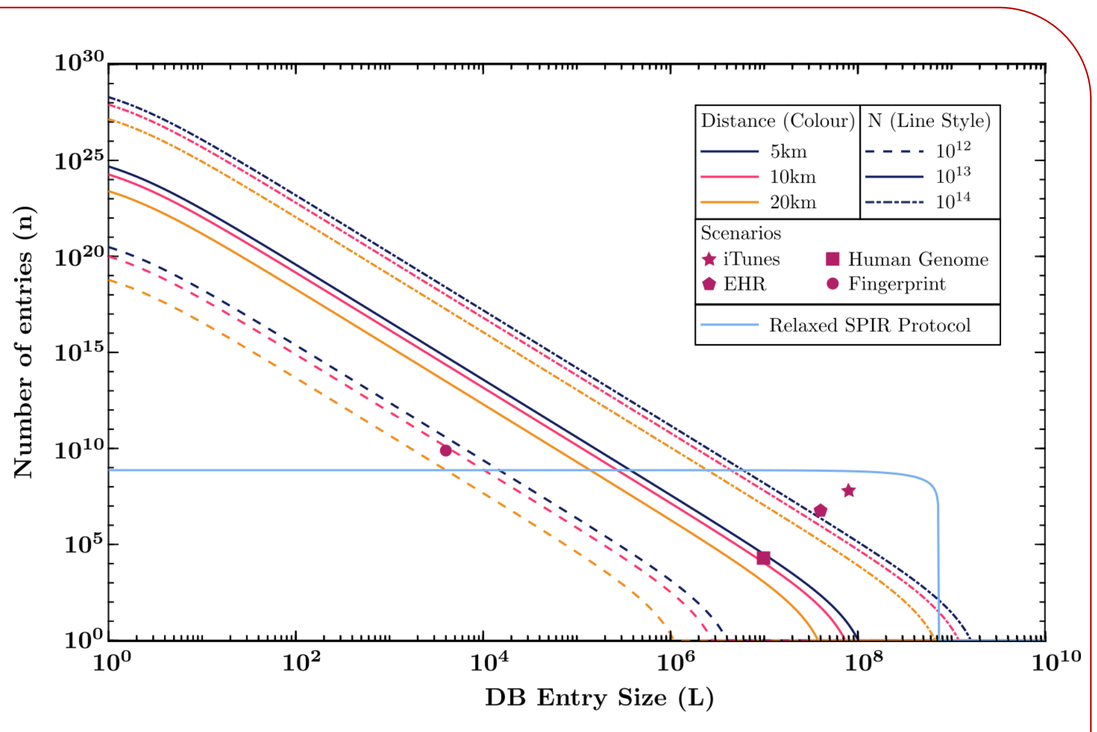

Figure 2: Plot of database parameters, L, the size of each entry of the database, and n, the number of entries in the database for various number of signals sent, N, (labelled by different line style) and distances, d (labelled by different colours). Four points are included that represents the database parameters of the usage scenarios described in the main text. Also included is a plot for a protocol that requires a more relaxed SPIR definition.

Discussion

In replacing classical secure channels with QKD, we introduced three additional assumptions: (1) the data centres do not intentionally leak the QKD keys to other parties including Eve, (2) all messages sent through the channels must be encrypted with OTP, and (3) data centres cannot access the classical channels used to establish the QKD keys after key exchange.

1. B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, J. ACM **45**, 965-981 (1998)
2. Y. Gertner, Y. Ishai, E. Kushilevitz, and T. Malkin, J. Comput. Syst. Sci. **60**, 592 (2000)
3. N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002)
4. M. Curty, F. Xu, C.C.W. Lim, K. Tamaki, and H.-K. Lo, Nat. Commun. **5**, 3732 (2014)
5. *Healthcare Broadband in America*, OBI Technical Paper 5 (Federal Communications Commission, Washington DC, USA, 2010)
6. *Information technology – Biometric data interchange formats – Part 2: Finger minutiae data*, ISO/IEC 19794-2:2011 (International Organization for Standardization, Geneva, Switzerland, 2011)
7. A. Piovesan, F. Antonaros, L. Vitale, P. Strippoli, M.C. Pelleri, and M. Caracausi, BMC Re. Notes **12**, 315 (2019)