論文の内容の要旨
Thesis Summary

論文題目: Design of Computationally Efficient Private Information Retrieval Protocols
（効率的な秘匿情報検索法の提案）

氏名: トマ　フランシス　ヴァネ

A key element to the widespread adoption of electronic communication was its strong guarantees of privacy and security. Traditional Public Key Encryption (PKE) systems developed from the 1970s fulfill those requirements. Indeed, any two entities can rely on such schemes to communicate safely and privately under strong computational assumptions. So strong in fact that many countries attempted to legally restrict its usage over concerns of national security. These regulations however proved hard to enforce and truly private communication between clients and remote services became the norm as the Internet developed. However, new privacy concerns stemmed from these omnipresent relations in mainly two ways. Some services built their business model around the systematic gathering of private information to create user profiles and the selling of said profiles to advertisers. Other services unwillingly shared this private data either because they were legally bound to do it or because a third party illegally breached their security and unauthorized access occurred.

One way to address these issues is Private Information Retrieval (PIR). As the name implies, PIR protocols allow a client to recover data on a server he has access to without compromising his privacy. The concept was formally introduced in 1995 by Chor *et al.* and a number of solutions were quickly designed. Compared to other cryptographic primitives, the definition of PIR only offers limited power to the client. Namely, he can only recover bits of information already present on the database and already available to him if he requests them directly. This apparent simplicity makes PIR comparatively likely to be used on real-world systems in the near future. Meanwhile, this low degree of freedom is still sufficient to be applied in a variety of scenarios ranging from financial services to biometric authentication.

There is a very natural so called trivial PIR algorithm satisfying all the privacy requirements and consisting in merely sending all the data stored on the server to the client regardless of his query. While not practical, this algorithm has a reasonable computational cost, optimal in a lot of settings. Its bottleneck resides in the prohibitive communication cost since network speeds tend to

be far slower than local data processing. As such, most early contributions to this research area focused on reducing the amount of data that has to be transferred and proved to be very successful in that regard.

More recently, focus has shifted towards the design of lightweight schemes as the heavy computation cost became the last obstacle standing between these theoretical protocols and practicality. My research takes place in this setting. Two main issues were identified and improved on. First, most schemes deal with a low-level model where the clients know the physical address of the bits he wishes to recover. While sufficient, this model can be highly impractical. Second, almost every known PIR protocol operates by reading the entire database for every query received. This is a major limitation which can be improved upon.

A scheme that allows the execution of SQL-like queries on a server-hosted database efficiently and privately was designed. Since most data publicly available online is organized in such databases, this is a necessary consideration to provide realistic schemes. This situates itself in a generalized PIR framework called Extended Private Information Retrieval. The algorithms described distinguish themselves from others in the literature since they do not require independent data replication, also known as multi-server PIR. This latter setting only makes sense in some very specific instances while our approach is applicable to a wide range of situations.

Another contribution is a lightweight block PIR scheme performing very efficiently. It is built by combining different existing constructions and its security is based on a well trusted computational assumption known as the Approximate GCD assumption. Straightforward preprocessing methods speeding up the running time are detailed and shown to also apply to other published protocols.

A new and more complex approach to PIR with preprocessing is created under the name Partial Server Side Parameter Selection (PSSPS) PIR. It achieves lower overall computation and allows for a justified trade-off where the cost of running the algorithm can be split in many ways between client and server. This approach is applied to the Approximate GCD-based scheme introduced previously and to another one from the literature. Further developments of this approach are discussed along with guidelines to create more efficient schemes based on it.

Now, all of the methods described above are compatible with each other and can be used together to build an efficient scheme on real-world databases or used independently and combined with other PIR schemes to gain different combinations of desirable properties.