個人情報保護に対応したデータ照合システムの開発

三上 春夫* 高山 喜美子 田中 留美

1. はじめに

罹患調査等の疫学調査においてデータベース間の照合は必須の作業である。これまで大規模なデータベース間のデータ照合の際には、照合するコンピュータまでデータを安全に搬送し、また持ち帰ることに細心の注意を払う必要があった。さらに照合時には一致した結果データのみならず本来閲覧されるべきでない他の不一致データを相手方に知られてしまう可能性があり、個人情報など守秘性の高いデータの扱いについては特段の配慮と手続が必要であった。

今回このようなデータベース間のデータ照合における諸問題を解決し、安全に照合結果を得るために「公開鍵暗号方式」を用いた照合システム Secure Data Matching System を開発することとした。

2. 公開鍵暗号方式

公開鍵暗号方式はデータの暗号化と復号化に一組の異なる鍵（鍵ペア）を用いる暗号化方式である。鍵ペアの一方で暗号化されたデータは対となるもう一方の鍵でしか復号できない。ユーザーA は鍵ペアを生成し、その一方（公開鍵）をメールやネット上で公開するとともに、もう一方の鍵（秘密鍵）を秘匿する。ユーザーB は A の公開鍵で自身のデータを暗号化し A に送信する。このデータを復号できるのは秘密鍵を所有する A に限られる（図 1）。

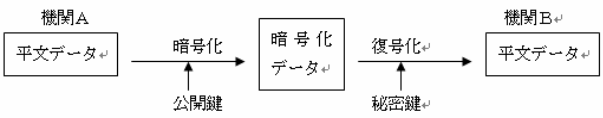


図 1. 公開鍵暗号方式の原理

3. 照合システム

以上の原理を照合システムに応用する（図 2）。

Step 1) 照合を実行する PC 上のシステムは最初に鍵ペアを生成し、公開鍵をユーザーに発行するとともに秘密鍵を内部に秘匿する。

Step 2) 各ユーザーは発行された公開鍵により原データを暗号化する。

Step 3) 同時に各ユーザーは持ち帰り用の再暗号化に使う公開鍵を各自のシステムにより生

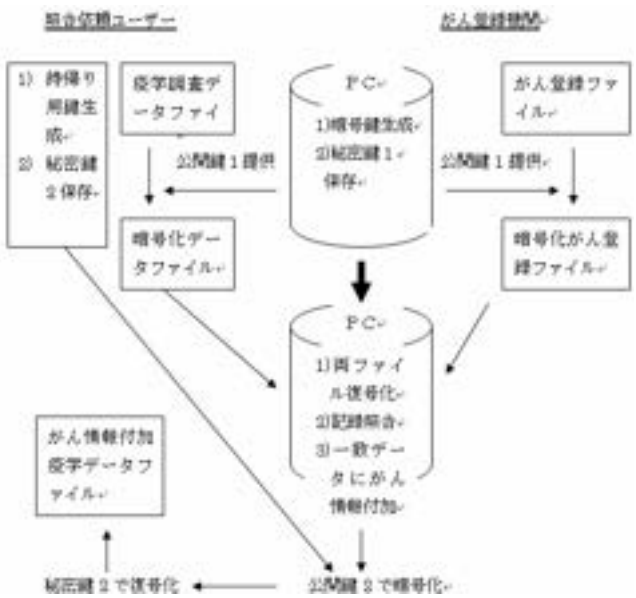


図 2. 公開鍵暗号方式を用いた匿名化照合システムの原理

*千葉県がんセンター疫学研究部
〒260-0801 千葉市中央区仁戸名町 666-2

成する。

Step 4) 照合データが入ると本システムは自身の秘密鍵で原データを復号後、照合作業を実行する。

Step 5) 本システムはユーザーに持ち帰り用の公開鍵の入力を求め、照合結果を再暗号化する。

Step 6) ユーザーは持ち帰った暗号化ファイルを各自システムの秘密鍵で復号し照合結果を得る。

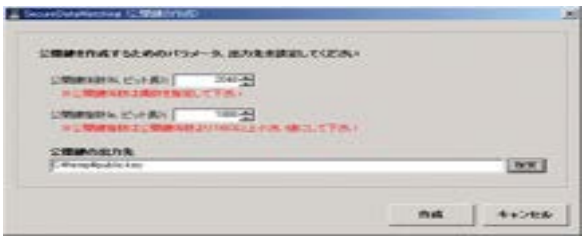
セキュリティ上の配慮から上記 Step４において、照合原データの復号、照合用インデックスの生成、照合時作業データの作成等は全てメモリ上に展開され、ハードディスク等の外部記憶装置上にファイルを作ることのない仕様とした。

照合は実行時に 2 つのファイルから対応させる項目どうしを複数組指定する。項目組は完全一致項目と不一致を許す項目組の指定が可能である。少なくとも 1 組の完全一致項目組を指定する必要がある。出力レコードの先頭には本システムにより項目が付加され、1)不一致の項目名、2)不一致の項目数、3)不一致の有無の 3 つのパターンのいずれかで一致の程度が報告される。

本システムでは商用システムにおいて公開鍵暗号方式の暗号化と復号化に広く用いられている RSA 方式を採用した。本方式の暗号／復号化ルーチンは㈱日立製作所が提供する Keymate/Crypto(R)により開発された。

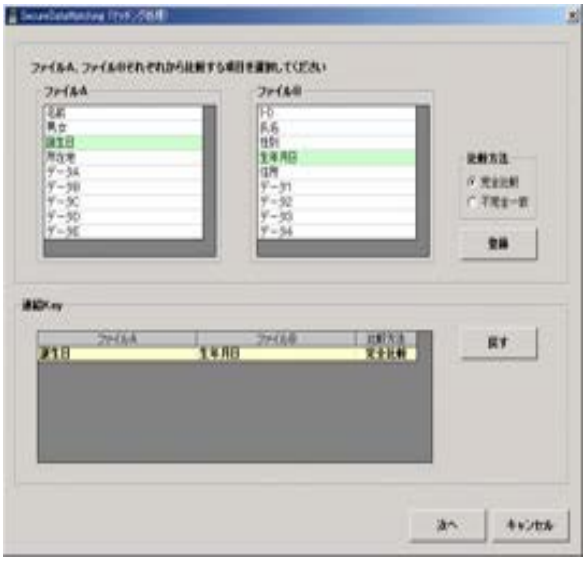
●鍵ペア生成画面

→公開鍵をファイルに書き込むと同時に秘密鍵を内部に暗号化して格納する。



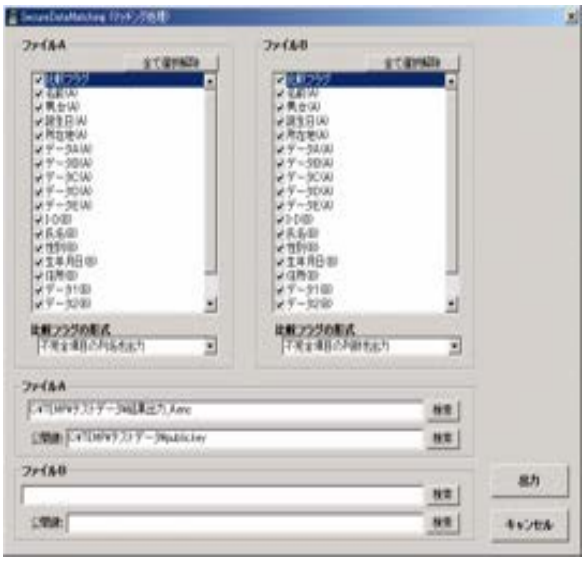
●マッチング項目の指定画面

→2 つのファイルの項目リストから対応する照合項目と完全一致か不一致を許すかを指定する。



●マッチング結果の出力指定画面

→持ち帰りのための公開鍵による暗号化が選択できる。持ち帰りの必要がない側はファイル名を指定しない。



●マッチング結果を復号化したファイル