部分復元可能な秘密分散バックアップした医療データ検索方法

1180362　沼　尚樹　【 ネットワーク信号処理研究室 】

1　はじめに

災害時に被災地域外からきた医師に対して，バックアップした医療データを提供することができれば，円滑に治療が行える．しかし，医療データは個人情報であるため治療に必要な情報のみを提供できればよい．そこで，部分復元可能な秘密分散法アルゴリズムが提案された [1]．災害時にバックアップしたデータを利用するには，患者のデータを検索しなければならない．しかし，部分復元可能な秘密分散を行ったデータを検索する仕組みは考えられておらず，検索するには，名前部分などを部分的に復元しなければならない．しかし，部分復元には時間がかかってしまい，素早く検索をすることができない．本研究では，部分復元可能な秘密分散法を行った医療データに検索用のシェアを付与し検索を行える仕組みを提案する．そのため，付与する情報の設定を行った．また，検索を素早く行うために付与する情報サイズの削減を行った．

2　提案手法

本研究では，部分復元可能な秘密分散法を行った医療データを検索する方法を提案する．提案手法はバックアップする医療データに，検索用の情報を付与し，検索を行えるようにする．図 1 に提案方法の流れを示す．バックアップを行う医療データに対し，そのデータの一部の情報を検索用の情報として利用する．検索用の情報は，分散し検索用シェアとして医療データに付与する．検索時には検索用のシェアを復元し，該当するデータを見つける．

3　付与する検索情報と優先度

検索情報として利用する情報とその優先度を表 1 に示す．これらの情報は全て医療データから得られる情報である．データを特定するには，検索情報を組み合わせて検索する必要がある．

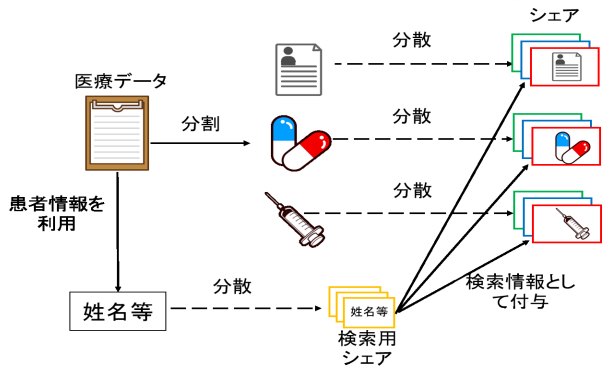

図 1　検索用情報の付与の流れ

表 1　検索情報とその優先度

検索情報	患者姓	患者名	性別	西暦	月日	住所	電話番号
優先度	1	4	2	3	6	7	5

表 2　3 文字以上の姓を検索した場合の中央値と最大人数

付与文字数	1 文字付与	2 文字付与	3 文字付与	4 文字付与
中央値 (人)	29517	12939	8068	8068
最大人数 (人)	117776	61969	54689	54689

患者データを検索するには，患者の情報を聞き出さなければならないが，災害時は患者自身から聞き出せない可能性がある．そのため，検索情報の聞き出しやすさが異なる．そこで，災害時にどのような順番で検索するのか検索情報の優先度を設定した．

3.1　姓情報の文字数の削減

災害時は素早く患者データを検索できなければならない．しかし，検索情報のサイズが大きければ検索に時間がかかってしまう．そのため，検索情報のサイズを削減する．本研究では，優先度の高い患者姓について文字数の削減を行う．日本人の姓と名の分布 [2] から東邦生命の被保険者と保険者の上位 60 位の姓の読みを利用し，先頭何文字でどのぐらいの人数 (744849 人) が当てはまるのかを調べた．表 2 に 3 文字以上の姓を検索した場合の中央値とヒットする最大の人数を示す．4 文字の姓も 3 文字あれば判断できることから姓については，3 文字付与する．

4　まとめ

本研究では，部分復元可能な秘密分散法を行った医療データ検索方法を提案した．また，付与する情報と災害時におけるその優先度を定めた．そして，素早く検索を行えるようにするために，検索情報のサイズの削減を提案した．姓情報について付与する文字数の削減を行った．今後は，他の検索情報についても検討する必要がある．

参考文献

[1] 田中麻実，福冨英次，福本昌弘，“秘密分散バックアップした医療データの部分復元,” 信学技報 IA2015-74，p.31-36，Dec.2015.

[2] 田中康仁，“日本人の姓と名の分布,” 日本オペレーションズ・リサーチ学会，Jun.1978.