令和 2 年度年度

学士学位論文

検索システムの構成による医療データ不正復元の防止

Prevention of Injustice Decoding of Medical Record by Configuring A Retrieval System

1210390　　吉冨　亮平

指導教員　　福本　昌弘

2021 年 3 月 5 日

高知工科大学 情報学群

要 旨

検索システムの構成による医療データ不正復元の防止

吉冨　亮平

広域災害に備えて医療データを保全し，被災地での診療に活用することが求められている．そこで，部分復元可能な秘密分散法が提案された．しかし，分散した医療データを検索する仕組みはない．そのため，分散した医療データを素早く検索する方法が提案された．検索方法は，医療データのシェアに個人を識別する情報を対応付けて保存してあり，ある条件において識別情報が絞られる．医療データの不正復元には，識別情報を一意に定め，医療データのシェアの組み合わせを知り，しきい値以上のシェアを集める必要があることから識別情報の絞り込みは大した問題ではない．しかし，個人情報である医療データを扱うため，識別情報の絞り込みが，医療データの不正復元に繋がるようなことがあってはならない．

本論文では，検索方法によって生じた識別情報の絞り込みの防止を目的として，識別情報が絞り込まれる要因となるデータの漏えいを防止する検索システムの構成を提案している．検索システム構成では，検索・復元を行う端末をリファレンスモニタにすることで医師端末の盗難による漏えいを防止する．また，ストレージ群とリファレンスモニタに対するアクセスを制限することで制限した端末以外からのアクセスによる漏えいを防止する．検索システム構成を用いることで安全に医療データの検索方法を使用することができる．

キーワード　秘密分散法，検索，検索システム構成

Abstract

Prevention of Injustice Decoding of Medical Record by Configuring A Retrieval System

Ryohei YOSHITOMI

It is required to preserve medical record in preparation for wide-area disasters and utilize it for medical treatment in the disaster area. Therefore, confined decodable secret sharing scheme was proposed. However, there is no mechanism for retrieving distributed medical record. Therefore, a method for quickly retrieving distributed medical record has been proposed. The retrieve method stores information that identifies an individual in association with the share of medical record, and the identification information is narrowed down under certain conditions. In order to injustice decode medical record, it is necessary to uniquely determine the identification information, know the combination of medical record shares, and collect the shares above the threshold value, so narrowing down the identification information is not a big problem. However, since medical record, which is personal information, is handled, narrowing down the identification information should not lead to unauthorized decoding of medical record.

In this paper, for the purpose of preventing the narrowing down of the identification information caused by the retrieval method, we have proposed the configuration of the retrieval system that prevents the leakage of data that causes the narrowing down of the identification information. In the retrieval system configuration, the terminal for retrieving and decoding is used as the reference monitor to prevent leakage due to theft of the doctor's terminal. In addition, by restricting access to the storage group

and reference monitor, leakage due to access from terminals other than the restricted terminal is prevented. By using the retrieval system configuration, it is possible to safely use the medical record retrieval method.

key words secret sharing scheme, retrieval, configureing a retrieval system

目次

目次

図目次

表目次

第 1 章

はじめに

1.1　本研究の背景と目的

東日本大震災の津波によって，沿岸部の病院に保管されていた電子カルテを含む医療データが消失し，被災地での医療行為に支障をきたした．これをうけて広域災害に備え医療データを保全し，被災地での診療に活用することが求められている [1]．災害時は，ネットワークや電源等のリソース不足が想定されるため，保全した医療データ全てを通信するには不安が残る．また，災害時の医療行為として投薬歴，アレルギー情報などの診療に必要な最低限のデータがあれば必ずしも全ての医療データが必要とは限らない．そこで，秘密分散法を用いた分散バックアップによる保全と欲しい情報のみを部分的に復元できる部分復元可能な秘密分散法が提案された [2]．医師が診療を行うには，分散バックアップした医療データから該当する患者の医療データ探す必要がある．しかし，提案された部分復元可能な秘密分散法は，分散した医療データを検索する仕組みはない．また，災害時は速やかに患者対応する必要があり，検索するだけでなく，素早く検索する必要がある．そのため，分散バックアップした医療データを素早く検索する方法が提案された [3]．提案された検索方法は医療データのシェアに個人を識別する情報を対応付けて保存してあり，ある条件において識別情報が絞られる．医療データの不正復元には，識別情報を一意に定め，医療データのシェアの組み合わせを知り，しきい値以上のシェアを集める必要があることから識別情報の絞り込みは大した問題ではない．しかし，個人情報である医療データを扱うため，識別情報の絞り込みが，医療データの不正復元に繋がるようなことがあってはならない．

本研究では，検索方法によって生じた識別情報の絞り込みの防止を目的として，識別情報

が絞り込まれる要因となるデータの漏えいを防止する検索システムの構成を提案し，評価する．

1.2　本論文の構成

本節では，本論文の構成について述べる．2 章では，医療データの分散バックアップについて述べ，部分復元可能な秘密分散法について述べる．3 章では，分散した医療データの検索方法について述べ，ある条件における識別情報の絞り込みについて述べる．4 章では，検索システム構成について述べる．5 章では，まとめと今後の課題について述べる．

第 2 章

遠隔地バックアップによる医療データの保全と災害時の活用

東日本大震災の津波によって，沿岸部の病院に保管されていた電子カルテを含む医療データが消失し，被災地での医療行為に支障をきたした．これをうけて広域災害に備えて，電子カルテを遠隔地にバックアップする取り組みが行われている．また，バックアップした医療データを被災地での医療行為に活用することが求められている．災害時は，ネットワークや電源等のリソース不足が想定される．そのため，全ての医療データを通信するには不安が残る．また，災害時の医療行為として投薬歴，アレルギー情報などの診療に必要な最低限のデータがあれば必ずしも全ての医療データが必要とは限らない．そこで，秘密分散法を用いた分散バックアップによる保全と欲しい情報のみを部分的に復元する部分復元可能な秘密分散法が提案された．

本章では，バックアップの対象となる医療データについて述べ，広域災害に備えた医療データのバックアップの条件について述べる．次に，保全方法の手法の 1 つとして考えられる (k, n) しきい値秘密分散法について述べ，災害時の活用を目的とした部分復元可能な秘密分散法について述べる．

2.1 対象となる医療データ

医療データには HIS(Hospital Information System) データやレセプトデータなど様々な種類があるが，バックアップの対象となる電子カルテについて述べる．また，厚生労働省

電子的診療情報交換推進事業において開発された標準規格である SS-MIX(Standerdized Structured Medical Information eXchange) についても述べる．

2.1.1　電子カルテ

電子カルテは，医師が患者を診察した際に作成される診療録を電子化し，電子情報としてカルテを編集，管理するシステムや仕組みのことを指す．電子保存であるため，物理的な保管場所を必要とせずネットワーク経由でどこからでも閲覧することが出来るというメリットやバックアップを容易に行うことが出来る．しかし，電子媒体であるため不正アクセスによる情報の漏えい，改ざん，破壊等の危険に晒される可能性がある．そのため，電子保存の三原則である真正性，見読生，保存性を満たし，最低 5 年間は保存することが義務付けられている．電子保存の三原則である真正性，見読性，保存性についての詳細を以下に示す [4].

真正性

真正性とは，正当な人が記録し確認された情報に関し第三者から見て作成の責任の所在が明確であり，かつ，故意または過失による，虚偽入力，書き換え，消去，及び混同が防止されていることである．

見読性

見読生とは，電子カルテに保存された内容を権限保有者からの要求に基づいて，診察等の活用に支障がない応答時間や肉眼で見読可能な状態にできることである．

保存性

保存性とは，記録された電子カルテが法的に定められた期間において，真正性，見読生を満たし保存することである．

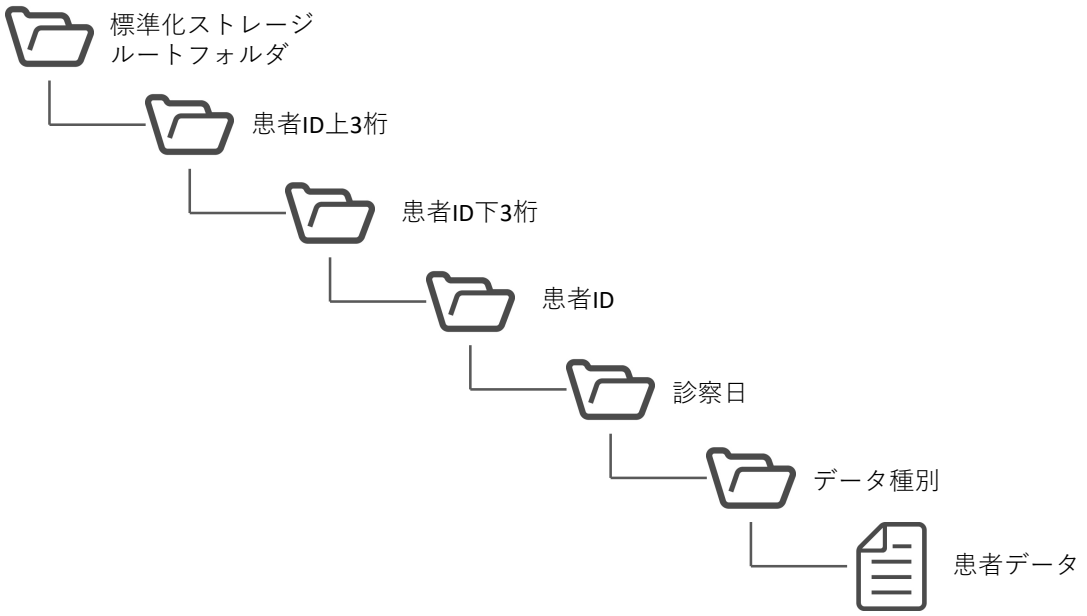

図 2.1　SS-MIX ディレクトリ構造

2.1.2　SS-MIX

各医療機関の電子カルテは様々なベンダーによって実装されているため，互換性がない．そのため，医療機関間での電子カルテを共有し地域医療連携することが困難であった．そこで，医療機関間の診療情報交換を推進するために電子カルテの標準化ストレージである SS-MIX が策定された [5]．標準化ストレージである SS-MIX はディレクトリ構造であり，ディレクトリ構造を図 2.1 に示す．SS-MIX のディレクトリ構造は，各医療機関用のルートフォルダーを置き，患者 ID によってディレクトリを分け配下に標準化された患者データを配置する．標準化された患者データを図 2.2 に示す．

標準化された患者データは標準規格である HL7(Health Level Seven) 形式である [5].

2.2　広域災害に備えた医療データのバックアップの条件

広域災害に備えた医療データのバックアップの条件について述べる．

広域災害に備えたバックアップとして，災害によってバックアップデータを保存している

```
MSH|^~\&|EGMAIN-GX|HIS|GW|GW|20141030195548.2901||OMP^O09|20141030195548659|P|2.5|
    |||||~ISO IR87||ISO 2022-1994
ZGW|9930000036|20141030195530|OMP-01^処方オーダ^L|14X303511256100|INS|01^総合診療科^L
SFT|Fujitsu Ltd.^D|V2|EGMAIN-GX
PID|0001||9930000036||ﾃｽﾄ^ﾃｽﾄ^^^^^L^I~ﾃｽﾄ^ﾃｽﾄ^^^^^L^P||19750505|M|||39|||||||||
    ||||||||N|||20130730150300|
ORC|NW|14X303511256100||1|||||20141030195536|KEN00100^検証^医師^^^^^^^L^^^^^I~ｹﾝｼｮｳ
    ^ｲｼ^^^^^^^L^^^^^P||KEN00100^検証^医師^^^^^^^L^^^^^I~ｹﾝｼｮｳ^ｲｼ^^^^^^^L^^^^^P
    |01^^^^^C||||01^総合診療科^L|MC20050D^^L|||||||||||O^外来患者オーダ^HL7-0482
TQ1|0001|1^日分|&&&&1日1回 起床時|||1^日分|2014103000
RXO|^タンナルビン末^MDCHOT^I1001390^タンナルビン末^L|1||G^ｸﾞﾗﾑ^MR9P^G^g^L
    |PWD^散剤､ドライシロップ剤^MR9P|TOC^OHP^MR9P^^外来処方~TOC^OHO^MR9P^^院外処方|||||1
    |G^ｸﾞﾗﾑ^MR9P^G^g^L||||Y|||||||1^G&ｸﾞﾗﾑ&MR9P&G&g&L
RXR|PO^口^HL7-0162
```

図 2.2　標準化された患者データ

ストレージが被災した場合やバックアップストレージのトラブル等が発生しても医療データを復元することができるよう冗長化する必要がある．また，冗長化していた場合でもストレージ全てが近くに存在する場合，広域災害によって全てが被災する可能性がある．そのため，遠隔地に保存する必要がある．医療データは個人情報であるため，バックアップデータから情報を読み取られないように秘匿化する必要がある．電子媒体の場合は，先程述べた電子保存の三原則を満たす必要がある．広域災害に備えた医療データのバックアップの条件を以下にまとめる．

1. バックアップデータの冗長化
2. バックアップデータの遠隔地保存
3. バックアップデータの秘匿化
4. 電子保存の三原則

2.3　(k, n) しきい値秘密分散法

広域災害に備えた医療データのバックアップ条件を満たす手法の 1 つには，データの秘匿化と分散による冗長化を図ることのできる (k, n) しきい値秘密分散法がある．(k, n) しきい値秘密分散法はデータの分散と復元の 2 つの段階で構成される．分散段階では，データから n 個のシェアと呼ばれる無意味な情報を生成し，分散する．復元段階では，分散した

2.3 (k, n) しきい値秘密分散法

n 個のシェアの内，k 個以上集めることで復元が可能である．k 個未満のシェアからは元のデータを復元することができないため，データの秘匿化が可能であり，$n-k$ 個以下のシェアを失ったとしても復元することが可能であるため冗長化することも可能である．本研究の (k, n) しきい値秘密分散法は，HL7 形式の電子カルテの分散バックアップを行うものである．以下に (k, n) しきい値秘密分散法の手法を示す.

分散したいデータを S，素数を $p(S < p$ かつ $n < p)$，$\mathbb{Z}/p\mathbb{Z} - \{0\}$ 上の集合 X を

$$X = \{x_1, x_2, \cdots, x_n\}(i \neq j \text{ のとき } x_i \neq x_j)$$

とする．X より

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{k-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{k-1} \end{pmatrix} \pmod{p}$$

となるような vandermonde 行列 X を作成する．また，$\mathbb{Z}/p\mathbb{Z}$ 上の乱数の集合 R を

$$R = \{r_1, r_2, \cdots, r_{k-1}\}(r_{k-1} \neq 0)$$

とする．S と R から $k \times 1$ の行列 A

$$A = \begin{pmatrix} S \\ r_1 \\ r_2 \\ \vdots \\ r_{k-1} \end{pmatrix}$$

を作成する．p を法とする X と A の乗算より

$$XA = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{k-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{k-1} \end{pmatrix} \begin{pmatrix} S \\ r_1 \\ r_2 \\ \vdots \\ r_{k-1} \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \pmod{p} \tag{2.1}$$

となるような $w_i(i = 1, 2, \cdots, n)$ が得られる．w_i をシェアと呼び，w_i と x_i を対応付けて分散バックアップする.

続いて，復元時の手順を示す．

復元時は，シェアを k 個集める．集めたシェアからシェア行列 W

$$W = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{pmatrix}$$

を作成する．選択したシェアに対応する x_i より

$$X' = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{k-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_k & x_k^2 & \dots & x_k^{k-1} \end{pmatrix} \pmod{p}$$

となるような $k \times k$ の vandermonde 行列 X' を作成する．X' の逆行列 X'^{-1} を W の左からかけると

$$\begin{aligned} X'^{-1}W &= X'^{-1} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{pmatrix} \\ &= X'^{-1}X'A \\ &= \begin{pmatrix} S \\ r_1 \\ r_2 \\ \vdots \\ r_{k-1} \end{pmatrix} \pmod{p} \end{aligned}$$

となり，データ S を復元することができる．以上が (k, n) しきい値秘密分散法の分散・復元

の手順である．式 (2.1) を連立方程式で表すと

$$\begin{cases} w_1 = S + r_1 x_1 + r_2 x_1^2 + \ldots + r_{k-1} x_1^{k-1} & (\mathrm{mod}\ p) \\ w_2 = S + r_1 x_2 + r_2 x_2^2 + \ldots + r_{k-1} x_2^{k-1} & (\mathrm{mod}\ p) \\ \qquad\qquad \vdots & \\ w_k = S + r_1 x_k + r_2 x_k^2 + \ldots + r_{k-1} x_k^{k-1} & (\mathrm{mod}\ p) \end{cases} \tag{2.2}$$

となる．式 (2.2) を解くことで S を求めることができる．w_i が k 個以上の場合は S を一意に定めることが出来るが，k 個未満の場合は S を一意に定めることができず，S を求めることができない．

以上のことから (k, n) しきい値秘密分散法は，k 個未満のシェアからは元のデータを復元することができないという秘匿性と $n - k$ 個以下のシェアを失ったとしても元のデータを復元することができるという冗長性を持っている．このことから個人情報である医療データのバックアップ手法として適している．

しかし，災害時の利用を考えたとき，ネットワークや電源等のリソース不足が想定されるため，全ての医療データを通信するには不安が残る．また，災害時の医療行為として投薬歴，アレルギー情報などの診療に必要な最低限のデータがあれば必ずしも全ての医療データが必要とは限らない．このことから欲しい情報のみを部分的に復元できるとよい．しかし，(k, n) しきい値秘密分散法はデータを復元するかしないかのどちらかであり，一部の情報だけを部分的に復元することはできない．

2.4　部分復元可能な秘密分散法

(k, n) 秘密分散法は，分散したデータを全て復元するかしないかのどちらかであり，一部の情報だけを部分的に復元することはできない．そこで，(k, n) しきい値秘密分散法を用いた部分復元アルゴリズムが提案された．本節では，部分復元アルゴリズムについて述べる．部分復元アルゴリズムは分散段階と復元段階で構成される．

2.4.1　分散段階

分散するデータ S を意味のある項目ごと d 個に分け，分けたデータを $S_i(i = 1, 2, \ldots, d)$ とする．このようにデータを意味のある項目ごとに分けることを分割と呼ぶ．分割したデータ S_i それぞれに対して (k, n) しきい値秘密分散法を用いてシェアを生成し，分散する．シェアの状態ではどのシェアがどの項目のシェアであるかの判別は困難であるため，項目が正しい順番で並んだデータを復元できるように，正しく紐付ける情報を作成し紐付け情報とする．紐付け情報は分割したデータの長さ l_i を要素として対応する項目ごとに並べたものである．秘匿したいデータの内容に対応する項目を 0 となるように作成する．

2.4.2　復元段階

分散したシェアから復元したいシェアのみをしきい値以上集め，秘匿したい項目を決めた紐付け情報を用いて，部分復元用シェアを作成する．部分復元用シェアに対して (k, n) しきい値秘密分散法と同様の復元操作を行うことで，復元したい項目のみを含んだデータを復元することができる．

2.4.3　部分復元の手順

秘密分散するデータ S を意味のある項目ごとに分割し，分割したデータを $S_i(i = 1, 2, \ldots, d)$，分割データサイズを l_i とする．データ S は分割したデータ S_i を用いて

$$S = S_1 \Pi_{m=2}^{d} 2^{l_m} + S_2 \Pi_{m=3}^{d} 2^{l_m} + \ldots + S_{d-1} 2^{l_d} + S_d \tag{2.3}$$

と表すことができる．分割したデータ S_i の集合を

$$S_{all} = \{S_1, S_2, \ldots, S_d\}$$

とおき，復元したいデータの集合を S_{all} の部分集合として

$$S_{all} \supseteq S_c = \{S_{c_1}, S_{c_2}, \ldots, S_{c_j}\} (1 \leqq j \leqq d)$$

としたとき，復元したいデータ S' は

$$S' = S_{c_1}\Pi_{m=2}^{j}2^{l_m} + S_{c_2}\Pi_{m=3}^{j}2^{l_m} + \ldots + S_{c_j}$$

と表すことができる．S_c の任意の要素である $S_{c_t}(1 \leqq t \leqq j)$ のシェア集合を

$$W_{c_t} = \{w_{c_{t,1}}, w_{c_{t,2}}, \ldots, w_{c_{t,n}}\}(1 \leqq t \leqq j)$$

とおく．復元したいデータのシェア集合 W_{c_t} を集めたものをシェア集合 G_c

$$G_c = \{W_{c_1}, W_{c_2}, \ldots, W_{c_j}\}$$

とおき，シェア集合 G_c の要素 W_{c_t} を部分復元シェアとなるように計算するための紐付け情報 u_c

$$u_c = \begin{pmatrix} \Pi_{m=2}^{j}2^{l_m} \\ \Pi_{m=3}^{j}2^{l_m} \\ \vdots \\ 2^{l_j} \\ 1 \end{pmatrix} \tag{2.4}$$

を作成する．部分復元の詳細な手順を以下に示す．

1. S を式 (2.3) を用いて分割し，$S_i(i = 1, 2, \ldots, d)$ を作成する．S_i のデータサイズを l_i とする．
2. 素数 $p(S<p$ かつ $n<p)$ を選択する．
3. $\mathbb{Z}/p\mathbb{Z} - \{0\}$ の集合 $X = \{x_1, x_2, \ldots, x_n\}\{i \neq j$ のとき $x_i \neq x_j\}$ から

$$X = \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^{k-1} \\ 1 & x_2 & x_2^2 & \ldots & x_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^{k-1} \end{pmatrix} \pmod{p}$$

となるような $n \times k$ の vandermonde 行列 X を作成する．また，分割データ S_i と $\mathbb{Z}/p\mathbb{Z}$ の集合 $R_i = \{r_{i,1}, r_{i,2}, \ldots, r_{i,k-1}\}$(ただし,$r_{i,k-1} \in \mathbb{Z}/p\mathbb{Z} - \{0\}$) からランダムに選択

し，行列 A_i

$$A_i = \begin{pmatrix} S_i \\ r_{i,1} \\ r_{i,2} \\ \vdots \\ r_{i,k-1} \end{pmatrix}$$

を作成する．

4. p を法とした X と X_i の乗算より

$$XA_i = \begin{pmatrix} w_{i,1} \\ w_{i,2} \\ \vdots \\ w_{i,n} \end{pmatrix} \pmod{p}$$

となるような $w_{i,j}(j = 1, 2, \ldots, n)$ をシェアと呼ぶ．$w_{i,j}$ を分散する．

5. G_c から各 w_{c_j} のシェアが k 個になるように集める．k 個のシェアから行列 W_c

$$W_c = \begin{pmatrix} w_{c_{1,1}} & w_{c_{2,1}} & \ldots & w_{c_{j,1}} \\ w_{c_{1,2}} & w_{c_{2,2}} & \ldots & w_{c_{j,2}} \\ \vdots & \vdots & \ddots & \vdots \\ w_{c_{1,k}} & w_{c_{2,k}} & \ldots & w_{c_{j,k}} \end{pmatrix}$$

を作成する．

6. W_c に対して式 (2.4) をかけ，部分復元用シェア $W_c u_c$ を作成する．

$$W_c u_c = \begin{pmatrix} w_{c_{1,1}} & w_{c_{2,1}} & \ldots & w_{c_{j,1}} \\ w_{c_{1,2}} & w_{c_{2,2}} & \ldots & w_{c_{j,2}} \\ \vdots & \vdots & \ddots & \vdots \\ w_{c_{1,k}} & w_{c_{2,k}} & \ldots & w_{c_{j,k}} \end{pmatrix} \begin{pmatrix} \Pi_{m=2}^{j} 2^{l_m} \\ \Pi_{m=3}^{j} 2^{l_m} \\ \vdots \\ 2^{l_j} \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} w_{c_{1,1}} \Pi_{m=2}^{j} 2^{l_m} + w_{c_{2,1}} \Pi_{m=3}^{j} 2^{l_m} + \ldots + w_{c_{j,1}} \\ w_{c_{1,2}} \Pi_{m=2}^{j} 2^{l_m} + w_{c_{2,2}} \Pi_{m=3}^{j} 2^{l_m} + \ldots + w_{c_{j,2}} \\ \vdots \\ w_{c_{1,k}} \Pi_{m=2}^{j} 2^{l_m} + w_{c_{2,k}} \Pi_{m=3}^{j} 2^{l_m} + \ldots + w_{c_{j,k}} \end{pmatrix} \quad (2.5)$$

式 (2.5) を展開してまとめると

$$W_c u_c = X \begin{pmatrix} S_{c_1}\Pi_{m=2}^{j} 2^{l_m} + S_{c_2}\Pi_{m=3}^{j} 2^{l_m} + \ldots + S_{c_j} \\ r_{1,1}\Pi_{m=2}^{j} 2^{l_m} + r_{2,1}\Pi_{m=3}^{j} 2^{l_m} + \ldots + r_{j,1} \\ \vdots \\ r_{1,k-1}\Pi_{m=2}^{j} 2^{l_m} + r_{2,k-1}\Pi_{m=3}^{j} 2^{l_m} + \ldots + r_{j,k-1} \end{pmatrix} \tag{2.6}$$

となる．

7. 式 (2.6) に X の逆行列 X^{-1} を左からかけると

$$X^{-1} W_c u_c = \begin{pmatrix} S_{c_1}\Pi_{m=2}^{j} 2^{l_m} + S_{c_2}\Pi_{m=3}^{j} 2^{l_m} + \ldots + S_{c_j} \\ r_{1,1}\Pi_{m=2}^{j} 2^{l_m} + r_{2,1}\Pi_{m=3}^{j} 2^{l_m} + \ldots + r_{j,1} \\ \vdots \\ r_{1,k-1}\Pi_{m=2}^{j} 2^{l_m} + r_{2,k-1}\Pi_{m=3}^{j} 2^{l_m} + \ldots + r_{j,k-1} \end{pmatrix}$$

となり

$$S' = S_{c_1}\Pi_{m=2}^{j} 2^{l_m} + S_{c_2}\Pi_{m=3}^{j} 2^{l_m} + \ldots + S_{c_j}$$

部分復元データ S' を復元できる．

2.5　まとめ

東日本大震災の津波によって，沿岸部の病院に保管されていた電子カルテを含む医療データが消失し，被災地での医療行為に支障をきたした．これをうけ広域災害に備えて，医療データの保全し，被災地での診療に活用することが求められている．本章では，バックアップの対象となるデータについて述べ，広域災害に備えた医療データのバックアップの条件について述べた．広域災害に備えた医療データのバックアップの条件から保全方法の 1 つとして (k, n) しきい値秘密分散法を用いた分散バックアップがある．(k, n) しきい値秘密分散法は，データの秘匿化と冗長化が可能だが，データを部分的に復元することはできない．そこで，(k, n) しきい値秘密分散法を用いた部分復元アルゴリズムが提案された．部分復元アルゴリズムにより，欲しい情報のみを部分的に復元することができるようになった．医師が患者の診療を行うには，分散バックアップした医療データから該当する患者の医療データを探

す必要がある．しかし，提案された部分復元可能な秘密分散法には，分散バックアップした医療データを検索する仕組みはない．

第 3 章

分散バックアップした医療データの検索方法と識別情報の絞り込み

広域災害に備えた医療データの保全と被災地での活用として，秘密分散法を用いた分散バックアップによる保全と欲しい情報のみを部分的に復元する部分復元可能な秘密分散法が提案された．医師が患者の診療を行うには，分散バックアップした医療データから該当する患者の医療データを探す必要がある．しかし，提案された部分復元可能な秘密分散法では，分散バックアップした医療データを検索する仕組みがない．考えられる方法として医療データのシェアの対応がわかるようにしてストレージに保存し，個人を識別できる情報を部分復元して検索する．しかし，復元は演算コストが高く，検索するのに時間がかかる．災害時は速やかに患者対応する必要があるため，素早い検索が必要である．そのため，この方法は災害時に適さない．そこで，分散バックアップした医療データを復元することなく，シェアの状態で検索する仕組みが提案された．提案された検索方法は，医療データのシェアに個人を識別できる情報を対応付けて保存している．しかし，ある条件において識別情報が有限個に絞り込まれる．

本章では，分散バックアップした医療データを復元することなく，シェアの状態で検索する仕組みについて述べる．また，ある条件における識別情報の絞り込みについて述べ，評価を述べる．

3.1　検索アルゴリズム

災害時，医師は患者の診療のためにバックアップデータから該当の患者のデータを検索し，診療を行う．しかし，提案された部分復元可能な秘密分散法では，検索の仕組みがない．そのため，名前などの個人を特定できる情報を部分的に復元してから検索する方法が考えられる．復元には連立方程式を解く必要があり，バックアップしている全ての人の識別情報を部分復元してから検索を行うため時間がかかることが想定される．災害時には通常より多くの患者が押し寄せることも想定されるため，素早く検索し診療する必要がある．よってこの方法は災害時に適していない．

検索の高速化を目的として，シェア間の係数の差を比較することでシェアの状態での検索を可能とする方法が提案された．この方法は $(2, n)$ しきい値秘密分散法を前提としている．また，提案方法は分散段階と検索段階で構成される．本節では，シェアの状態での検索を可能とする仕組みについて分散段階と検索段階に分けて述べる．

3.1.1　分散段階

分散バックアップしたい医療データから個人を特定できるキーワード $kw > 0$ を生成する．素数 $p(kw < p, n < p)$ とし，$\mathbb{Z}/p\mathbb{Z} - \{0\}$ 上の集合を X とする．また，p, X は医療データを秘密分散した際とは異なる値にする．X は

$$X = \{x_1, x_2, \ldots, x_n\} (i \neq j \text{ のとき } x_i \neq x_j)$$

とすることができ，X より

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \pmod{p}$$

となるような vandermonde 行列 X を作成する．kw と $\mathbb{Z}/p\mathbb{Z} - \{0\}$ 上の乱数 r より

$$A = \begin{pmatrix} kw \\ r \end{pmatrix}$$

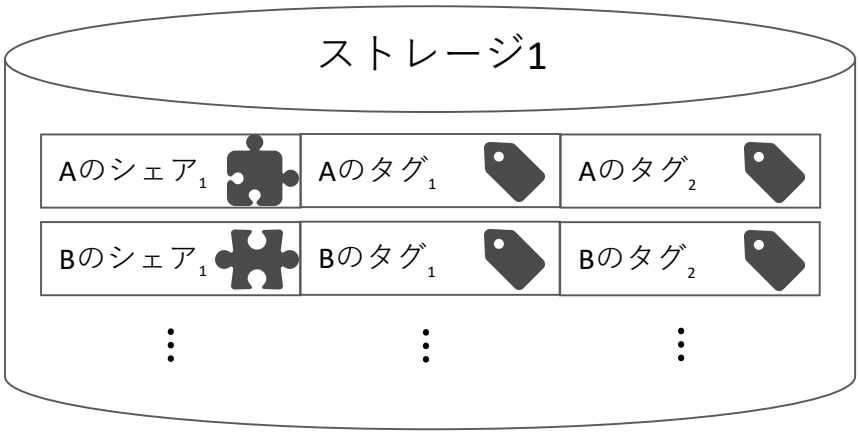

図 3.1　シェアとタグの保存方法

となるような行列 A を作成する．X と A の乗算より

$$XA = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} \pmod{p}$$

となるようなタグ $t_u (u = 1, 2, \ldots, n)$ が得られる．また，方程式で表すと

$$t_u \equiv kw + rx_u \pmod{p} \tag{3.1}$$

となる．医療データのシェアに 2 つずつ対応付けて分散バックアップする．タグと対応する医療データのシェアの保存方法を図 3.1 に示す．また，タグ作成に使用した行列 X と p は保存しておく必要がある．

3.1.2　検索段階

検索したいキーワードを kw' とする．$\mathbb{Z}/p\mathbb{Z} - \{0\}$ 上の乱数 r' と kw' より

$$A' = \begin{pmatrix} kw' \\ r' \end{pmatrix}$$

となるような行列 A' を作成する．分散段階で作成した行列 X と A' の乗算より

$$XA' = \begin{pmatrix} t'_1 \\ t'_2 \\ \vdots \\ t'_n \end{pmatrix}$$

となるような検索タグ t'_u が得られる．また，方程式で表すと

$$t'_u \equiv kw' + r'x_u \pmod{p} \tag{3.2}$$

となる．分散バックアップした医療データに付与されているタグ t_u と作成した検索タグ t'_u の減算を行う．式 (3.1), 式 (3.2) より

$$t_u - t'_u \equiv (kw - kw') + (r - r')x_u \pmod{p} \tag{3.3}$$

となる．式 (3.3) の両辺に x_u の逆元 x_u^{-1} を掛けると

$$(t_u - t'_u)x_u^{-1} \equiv (kw - kw')x_u^{-1} + (r - r') \pmod{p} \tag{3.4}$$

となる．

以上の手順より，シェアの状態での検索が可能となる．

3.2　識別情報の絞り込み

提案された検索方法では，タグ t_u は医療データのシェアと対応付けて 2 つずつ保存されているため，攻撃者が医療データのシェアにアクセスすることができれば，対応するタグ t_u は漏えいする．また，この検索方法はタグ t_u を生成する際のしきい値 k が 2 で固定であるため，この検索方法を使用されていることが知られていた場合，k は明らかとなる．p, x_u は全てのタグ t_u で共通であるため，タグ t_u に対応付けて保存する必要はなく，医師が使用する端末に保存しておくことが考えられる．しかし，災害時は病院外の仮設的な診療所で診療が行われ，医師端末はその場での使用が想定される．そのため，医師端末の盗難などにより p, x_u が漏えいする可能性がある．そこで t_u, k と p または x_u が漏えいした際の識別情報 kw の絞り込みについて述べる．

3.2.1　t_1, t_2, k, p が漏えいした場合

n 個のうちの t_1, t_2 と k, p が漏えいした場合について考える．t_1, t_2 を作成する際の条件より kw の範囲は $1 \leqq kw \leqq p-1$ である．p が漏えいした場合，kw は $p-1$ 通りに絞ることができる．また，t_1, t_2, k に加えて p が漏えいしている場合，kw を求めるためには，t_1, t_2 と対応する $x_i, x_j (i, j \in u, i \neq j)$ があれば良い．x_u も t_u を作成する際の条件より，$x_u \in \mathbb{Z}/p\mathbb{Z} - \{0\}, x_i \neq x_j$ であるため，x_i, x_j を有限個に絞ることができる．x_i, x_j を総当たりすることで kw を絞り込むことが可能である．

絞り込み手順

kw を絞り込む手順について述べる．

x_i, x_j の選び方は

$$ {}_{p-1}P_2 = (p-1)(p-2) \tag{3.5} $$

通りある．t_1, t_2 より

$$ T = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} $$

となるような行列 T を作成する．x_i, x_j より

$$ X = \begin{pmatrix} 1 & x_i \\ 1 & x_j \end{pmatrix} $$

となるような行列 X を作成する．行列 X の逆行列 X^{-1} は

$$ X^{-1} = \frac{1}{x_j - x_i} \begin{pmatrix} x_j & -x_i \\ -1 & 1 \end{pmatrix} $$

となり，X^{-1} を T の左から掛けると

$$ X^{-1}T = \frac{1}{x_j - x_i} \begin{pmatrix} x_j & -x_i \\ -1 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} $$

$$ = \begin{pmatrix} \frac{t_1 x_j - t_2 x_i}{x_j - x_i} \\ \frac{t_2 - t_1}{x_j - x_i} \end{pmatrix} $$

となり，kw は

$$kw \equiv \frac{t_1 x_j - t_2 x_i}{x_j - x_i} \pmod{p}$$

となる．このとき kw は $(p-2)$ 通り以下に絞られる．

演算量

計算コストの高い乗算演算に着目して演算量を求める．乗算回数を求める．

2×2 の行列 X の逆行列 X^{-1} を掃き出し法を用いて作成すると乗算回数は

$$2^3$$

回となる．x_i, x_j の選び方は式 (3.5) より，$(p-1)(p-2)$ 通りあるため逆行列 X^{-1} の乗算回数は

$$2^3(p-1)(p-2) \tag{3.6}$$

回となる．$X^{-1}T$ に必要な乗算回数は

$$2^2$$

回となる．x_i, x_j の選び方は式 (3.5) より，$(p-1)(p-2)$ 通りあるため逆行列 X^{-1} は $(p-1)(p-2)$ 通り作成され，$X^{-1}T$ の乗算回数は

$$2^2(p-1)(p-2) \tag{3.7}$$

回となる．合計乗算回数は式 (3.6), (3.7) より

$$2^3(p-1)(p-2) + 2^2(p-1)(p-2) = 12(p-1)(p-2) \tag{3.8}$$

回となる．

1 回の乗算演算にかかる演算量を d^2 とすると，演算量は

$$d^2\{12(p-1)(p-2)\} \tag{3.9}$$

となる．

表 3.1　実験環境

OS	Ubuntu20.04.1LTS
CPU	Intel(R)Xeon(R)W-1290P CPU@3.70GHZ
メモリ	16GB
コンパイラ	gcc–9.3
コンパイラオプション	-lgmp

評価

識別情報の絞り込みについて絞り込まれる範囲と演算量から評価する．

識別情報の絞り込み手順から kw は $(p-2)$ 通り以下に絞られる．そのため，p を大きくすることで，絞り込まれる kw の範囲は広がる．t_1, t_2, p を 64bit で作成した場合，64bit 以下の最大の素数 18446744073709551557 を p とする．表 3.1 の実験環境において 64bit 同士の乗算にかかる時間は $0.0000001884(s)$ であった．よって式 (3.9) の演算量より絞り込みにかかる時間は約 $7.69 \times 10^{32}(s)$ であった．p を大きくすることで，絞り込まれる範囲が広がることや識別情報を絞り込む時間がさらにかかることから識別情報の絞り込みは有効な手段ではないと考えられる．そのため，識別情報の絞り込みは大した問題ではない．

3.2.2　t_1, t_2, k, x_1, x_2 が漏えいした場合

n 個のうちの t_1, t_2 と対応する x_1, x_2 が漏えいした場合，連立方程式を解くことができるが，剰余演算の法 p がわからないため，正しい kw を求めることができない．そこで，除算を行わず，連立方程式を解き進めた結果に p を総当たりすることで kw を有限個に絞ることが可能である．

絞り込み手順

kw を絞り込む手順について述べる．

t_1, t_2 より

$$T = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$$

となるような行列 T を作成する．t_1, t_2 に対する x_1, x_2 より

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix}$$

となるような行列 X を作成する．行列 X の逆行列 X^{-1} は

$$X^{-1} = \frac{1}{x_2 - x_1} \begin{pmatrix} x_2 & -x_1 \\ -1 & 1 \end{pmatrix}$$

となり，X^{-1} を T の左からかけると

$$X^{-1}T = \frac{1}{x_2 - x_1} \begin{pmatrix} x_2 & -x_1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{t_1 x_2 - t_2 x_1}{x_2 - x_1} \\ \frac{t_2 - t_1}{x_2 - x_1} \end{pmatrix}$$

となる．除算を行わずに計算した結果の $(1,1)$ 成分を KW とすると KW は

$$KW = \frac{t_1 x_2 - t_2 x_1}{x_2 - x_1} \tag{3.10}$$

となる．KW に対して p を法として除算した結果が kw となる．

式 (3.10) における KW が自然数に割り切れるとき KW を

$$KW = a \quad (a \in \mathbb{N}) \tag{3.11}$$

とすると $a = p$ となるとき $kw = 0$，$a < p$ になるとき kw は全て a になる．したがって kw は最大で a 未満の素数の個数 $+1$ 通りになり，絞ることができる．KW の取る最大値は $(p-1)^2$ になる．したがって kw は最大で $(p-1)^2$ 未満の素数の個数通りになり，絞ることができる．KW が自然数に割り切れる確率を図 3.2 に示す．p を大きくすることで，KW が自然数に割り切れる確率は減少する．

式 (3.10) において KW が自然数に割り切れないとき kw は不定である．

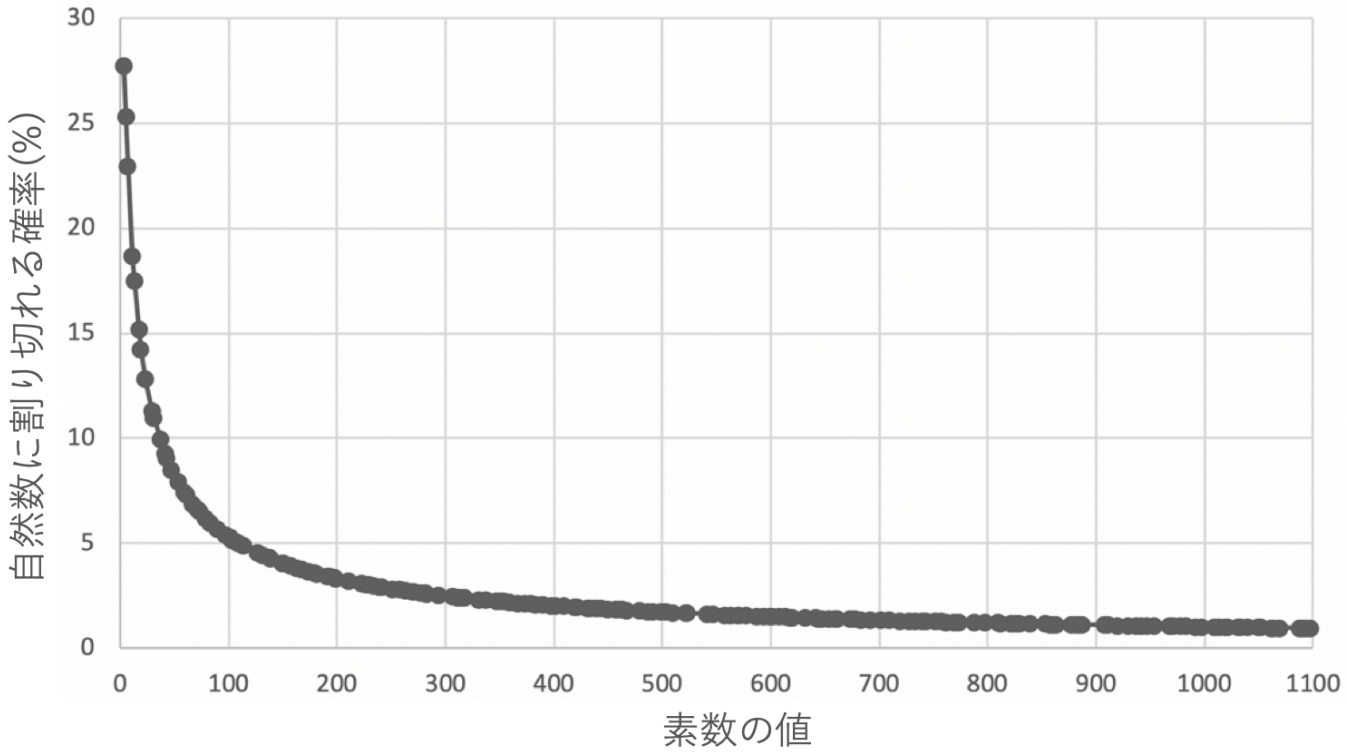

図 3.2　KW が自然数に割り切れる確率

演算量

t_1, t_2, k, x_1, x_2 が漏えいし，KW が自然数に割り切れる場合の演算量について述べる．計算コストの高い乗算・除算演算に着目して演算量を求める．

2×2 の行列 X の逆行列 X^{-1} を掃き出し法を用いて作成すると乗算回数は

$$2^3$$

回となる．$X^{-1}T$ に必要な乗算回数は

$$2^2$$

回となる．剰余演算の法 p による除算は行わないため，KW は式 (3.6) の分母 $(x_2 - x_1)$ の逆元 $(x_2 - x_1)^{-1}$ を用いると

$$KW = (t_1 x_2 - t_2 x_1)(x_2 - x_1)^{-1}$$

と表すことができる．

$(x_2 - x_1)$ の逆元 $(x_2 - x_1)^{-1}$ を求めるのに必要な乗算・除算回数について述べる．

式 (3.11) の a が素数のとき a 以下の素数の個数を b とし，素数を $\{p_1, p_2 \ldots, p_{b-1}, a\}$ とする．a が素数のときの $(x_2 - x_1)$ の逆元 $(x_2 - x_1)^{-1}$ の乗算・除算回数は

$$\frac{1}{2}\{\log(p_1) + \log(p_2) + \ldots + \log(p_{b-1}) + \log(a)\}$$

回となる．

式 (3.11) の a が素数でないとき a より大きい最初の素数を A とし，A 以下の素数の個数を c としたとき素数は $\{p_1, p_2 \ldots, p_{c-1}, p_A\}$ とする．a が素数でないときの $(x_2 - x_1)$ の逆元 $(x_2 - x_1)^{-1}$ の乗算・除算回数は

$$\frac{1}{2}\{\log(p_1) + \log(p_2) + \ldots + \log(p_{c-1}) + \log(A)\}$$

となる．a が素数のときの合計乗算回数は

$$12 + \frac{1}{2}\{\log(p_1) + \log(p_2) + \ldots + \log(p_{b-1}) + \log(a)\}$$

回となる．a が素数でないときの合計乗算回数は

$$12 + \frac{1}{2}\{\log(p_1) + \log(p_2) + \ldots + \log(p_{c-1}) + \log(A)\}$$

回となる．

1 回の乗算・除算演算にかかる演算量を d^2, D^2 とすると，演算量は a が素数のとき

$$d^2[12 + \frac{1}{2}\{\log(p_1) + \log(p_2) + \ldots + \log(p_{b-1}) + \log(a)\}]$$

$$+\frac{1}{2}D^2\{\log(p_1) + \log(p_2) + \ldots + \log(p_{b-1}) + \log(a)\}$$

となる．a が素数でないとき

$$d^2[12 + \frac{1}{2}\{\log(p_1) + \log(p_2) + \ldots + \log(p_{c-1}) + \log(A)\}]$$

$$+\frac{1}{2}D^2\{\log(p_1) + \log(p_2) + \ldots + \log(p_{c-1}) + \log(A)\}$$

となる．

評価

識別情報の絞り込みについて絞り込まれる範囲，絞り込みが可能となる確率や演算量から評価する．

絞り込み手順から kw は最大で $(p-1)^2$ 未満の素数の個数通りになる．そのため，p を大きくすることで，kw の範囲は広がる．また，識別情報の絞り込みが可能となるのは KW が自然数に割り切れるときであり，図 3.2 より p を大きくすることで KW が自然数に割り切れる確率は減少する．t_1, t_2, x_1, x_2 を 64bit で作成した場合，64bit 同士の乗算・除算にかかる時間は表 3.1 の実験環境よりそれぞれ $0.0000001884(s)$，$0.0000002405(s)$ であった．式 (3.10) の KW の取りうる最大値は $(p-1)^2$ になり，$(p-1)^2$ より大きい最初の素数を $p' = 340282366920938463426481119284349108409$ とする．p' 以下の素数の個数を約 3.835×10^{36} 個として，$(p-1)^2$ 未満の素数の $(x_2 - x_1)$ の逆元を求める乗算・除算回数を全て $\frac{1}{2}(\log(p'))$ とすると，約 $3.17 \times 10^{31}(s)$ であった．p を大きくすることで，kw の範囲は広がり，KW が自然数に割り切れる確率が減少することや識別情報を絞り込む時間がさらにかかることから識別情報の絞り込みは有効な手段ではないと考えられる．そのため，識別情報の絞り込みは大した問題ではない．

3.3　まとめ

医師が患者の診療を行うには，分散バックアップした医療データから該当する患者の医療データを探す必要がある．しかし，部分復元可能な秘密分散法において分散バックアップした医療データを検索する仕組みはない．また，災害時は素早く患者対応するため，素早い検索が必要となる．本章では，検索の高速化を目的としたシェア間での係数の差の比較によってシェアの状態で検索する仕組みについて述べ，検索方法を用いることで，ある条件において識別情報が有限個に絞られることについて述べた．識別情報の絞り込みは剰余演算の法を大きくすることや医療データの不正復元には，識別情報を一意に定め，医療データの組み合わせを知り，しきい値以上の医療データのシェアを集め，医療データのシェア作成に使用し

た p, X を用いる必要があることから大した問題ではない．しかし，医療データは個人情報であるため，識別情報の絞り込みが医療データの不正復元されるきっかけになるようなことがあってはならない．

第 4 章

検索システムの構成

提案された検索方法を使用した場合，医療データのシェアの組み合わせを知ることができる識別情報の復元に必要な t_u, p, x_u が漏えいする恐れがある．加えて，n 個のうち 2 つの t_u, k, p または n 個のうち 2 つの $t_u, k,$ と t_u に対応する 2 つの x_u が漏えいした場合，識別情報 kw が有限個に絞られる．しかし，識別情報の絞り込みは剰余演算の法を大きくすることや医療データの不正復元には，識別情報を一意に定め，医療データの組み合わせを知り，しきい値以上の医療データのシェアを集め，医療データのシェア作成に使用した p, X を用いる必要があることから大した問題ではない．識別情報の絞り込みは大した問題ではないが，医療データは個人情報であるため，識別情報の絞り込みが医療データの不正復元されるきっかけになるようなことがあってはならない．そのため，識別情報が絞り込まれる要因となる t_u, p, x_u の漏えいを防止する必要がある．

本章では，部分復元可能な秘密分散法と検索方法を考慮した検索システムによる t_u, p, x_u の漏えいを防止する．そのため，検索システムの構成要素について述べ，t_u, p, x_u の漏えいを防止する検索システム構成について述べ，検索システム構成について評価する．

4.1　検索システムの構成要素

検索システム構成を提案するために必要な部分復元可能な秘密分散法と検索方法で扱うデータと想定する利用者について述べ，検索システムで扱うデータと利用者から検索システム構成に必要なデバイスについて述べる．

表 4.1　検索システムで扱うデータ一覧

使用する段階	データの種類
分散段階	・医療データ　・シェア ・$x_u(med)$ ・$p(med)$ ・$r(med)$ ・識別情報　・タグ t_u ・$x_u(tag)$ ・$p(tag)$ ・$r(tag)$
検索段階	・検索キーワード　・検索タグ ・$x_u(tag)$ ・$p(tag)$ ・r' ・タグ t_u
復元段階	・医療データ ・シェア ・$x_u(med)$ ・$p(med)$ ・$r(med)$

4.1.1　検索システムで扱うデータ

部分復元可能な秘密分散法と検索方法から検索システムで扱うデータについて表 4.1 に示す.

検索システムで扱うデータについての説明を行う．データ形式 (テキスト，バイナリ) は問わず，意味データとする．ただし，秘密分散法を用いて作成したシェアは元データ (平文) とは区別するものとする.

検索システムでは，電子カルテを医療データとする．医療データを秘密分散して生成される無意味な情報をシェアとする．医療データを秘密分散する際に使用する集合 X を $x_u(med)$，素数 p を $p(med)$，乱数 r を $r(med)$ とする．電子カルテ内にある個人を特定できる情報を識別情報 kw とする．識別情報は氏名やマイナンバーが考えられる．識別情報を秘密分散して生成される無意味な情報をタグ t_u とする．医療データの検索時に入力する情報を検索キーワード kw' とする．検索キーワードを秘密分散して生成される無意味な情報を検索タグ t'_u とする．識別情報と検索キーワードを秘密分散する際に使用する共通の集合 X を $x_u(tag)$，共通の素数 p を $p(tag)$ とする．識別情報を秘密分散する際に使用する乱数 r を $r(tag)$ とし，検索キーワードを秘密分散する際に使用する乱数 r を r' とする．検索システムで扱うデータ説明についてに表 4.2 に示す.

表 4.2　検索システムで扱うデータ説明

データ名	データの説明
医療データ	標準化された患者の診療録 (電子カルテ)
シェア	医療データを秘密分散して生成される無意味な情報
$x_u(med)$	医療データを秘密分散する際に使用する集合 X
$p(med)$	医療データを秘密分散する際に使用する素数 p
$r(med)$	医療データを秘密分散する際に使用する乱数 r
識別情報 kw	電子カルテ内にある個人を特定できる情報 (氏名やマイナンバーを想定)
タグ t_u	識別情報を秘密分散して生成される無意味な情報
検索キーワード kw'	医療データの検索時に入力する情報
検索タグ t'_u	検索キーワードを秘密分散して生成される無意味な情報
$x_u(tag)$	識別情報と検索キーワードを秘密分散する際に使用する共通の集合 X
$p(tag)$	識別情報と検索キーワードを秘密分散する際に使用する共通の素数 p
$r(tag)$	識別情報を秘密分散する際に使用する乱数 r
r'	検索キーワードを秘密分散する際に使用する乱数 r

4.1.2　検索システムの利用者

検索システムを利用する人物は災害時に開設される仮設の診療所で診療する医師とする．医師以外の人物はその他の人物とする．その他の人物には悪意をもった攻撃者やシステムを利用する権限のない人物が含まれる．検索システムの利用権限について述べる．

医師が患者の診療を始めるには，分散バックアップした医療データから該当する患者を探す必要がある．そのため，検索システムで検索キーワードを入力する．検索にヒットした医療データを閲覧し，診療した結果を書き込む必要がある．その他の人物が検索システムを利用できた場合，医療データを検索し，閲覧することができてしまう．医療データは個人情報

表 4.3　検索システムの利用者権限

	検索キーワードの入力	医療データの閲覧	医療データの書き込み
医師	○	○	○
その他の人物	×	×	×

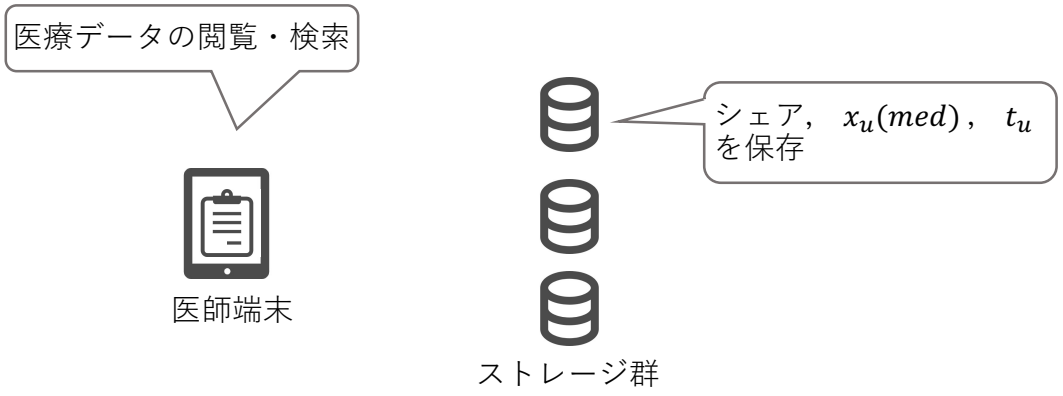

図 4.1　検索システムの構成デバイス

であるため，その他の人物が検索システムにおいて，いかなる操作もできないようにする必要がある．検索システムの利用者権限について表 4.3 に示す．

4.1.3　検索システムの構成デバイス

検索システムで扱うデータと利用者から医師が医療データを閲覧，書き込み，検索キーワードを入力する端末が必要となる．この端末を医師端末とする．次に，シェア等を遠隔地に分散バックアップするストレージ群が必要となる．検索システムの構成デバイスについて図 4.1 に示す．

4.2　検索システムの構成デバイスからの漏えいのリスク

識別情報が絞り込まれる要因となり，秘匿したい $t_u, p(tag), x_u(tag)$ の漏えいリスクについて述べる．

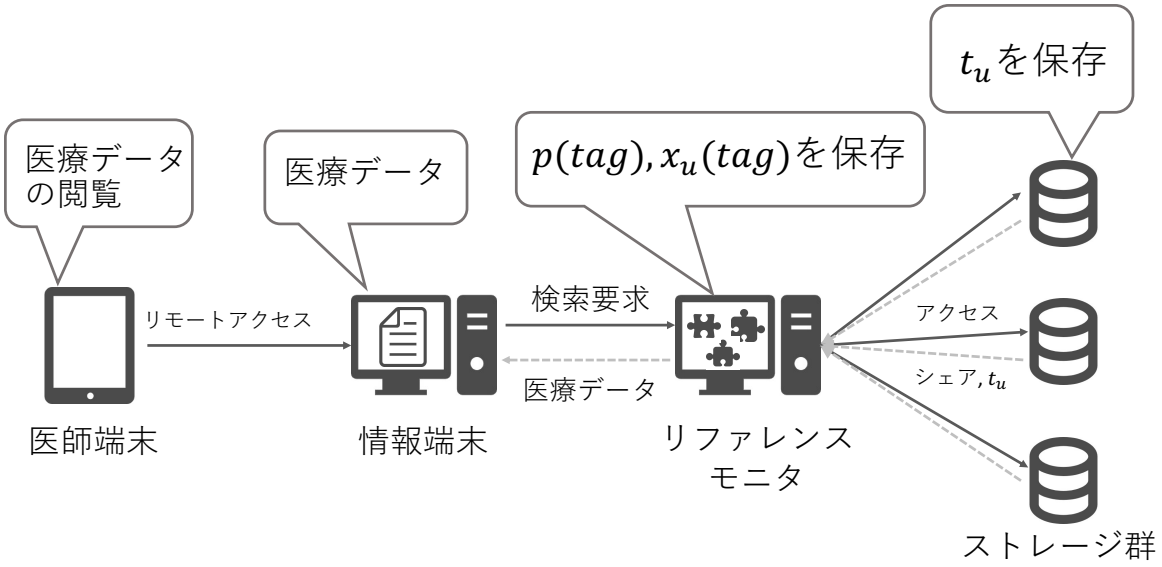

図 4.2　検索システム構成

t_u は医療データのシェアに対応付けてストレージに保存しているため，攻撃者がストレージにアクセスすることができれば，シェアに対応した t_u は漏えいする．$p(tag), x_u(tag)$ は全ての t_u で共通であるため，ストレージに保存する必要はないことから医療データの検索・復元を行う医師端末に保存されることが考えられる．医師端末は災害時に開設される仮設の診療所での利用が想定されるため，医師端末の安全性を確保することができないことから，盗難によって $p(tag), x_u(tag)$ が漏えいする恐れがある．

これらのことから，ストレージからの t_u の漏えいを防止するために，ストレージへのアクセスできる端末を制限する必要がある．また，$p(tag), x_u(tag)$ の漏えいを防止するために $p(tag), x_u(tag)$ へアクセスできる端末を制限し，盗難の恐れのない場所に保存する必要がある．

4.3　検索システム構成

検索システムの構成について述べ，医師端末，情報端末，リファレンスモニタについて述べ，各端末における流れや必要な照合について示す．

t_u が保存されているストレージへのアクセスを制限するために，ストレージへアクセスできる端末をリファレンスモニタと呼び，災害拠点病院内に設置する．病院内に設置することで安全は保証できるものとする．医師端末の盗難により $p(tag), x_u(tag)$ が漏えいする恐れがあるため，検索・復元を行う端末をリファレンスモニタとして $p(tag), x_u(tag)$ を保持する．

この場合，医師が診療する際は，医師端末がリファレンスモニタに検索要求を送ることで医療データを取得する．そのため，リファレンスモニタへ検索要求を送ることができれば医療データが漏えいする恐れがある．

そこで，リファレンスモニタへの検索要求を制限するために，検索要求を送信できる端末を情報端末と呼び，災害拠点病院内に設置する．病院内に設置することで安全は保証できるものとする．また，医師端末の安全は保証できないことからどのような情報も保持しないことが望ましい．そのため，医師が診療を行う際は，医師端末から情報端末へリモートアクセスすることで医師端末に医療データを保存することなく医療データの閲覧や検索が行える．検索システム構成を図 4.2 に示す．

4.3.1 医師端末

医師端末はどのような情報も保存しないことが望ましいことから，医師が診療する際は医師端末から情報端末へリモートアクセスすることで医師端末に医療データを保存することなく医療データの閲覧や検索を行う．医師端末から情報端末へリモートアクセスする流れを図 4.3 に示す．医師端末を医師に支給する際に医師端末の利用者が正当であるかを確認し，パスワードや生体認証の登録を行う．医師がリモートアクセスする際は，医師端末の利用者照合をクリアした場合，正当な利用者としてリモートアクセスが可能となる．リモートアクセスを行なうために必要や照合の様子を図 4.4 に示す．

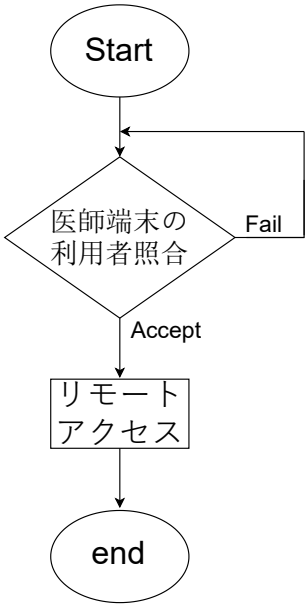

図 4.3　リモートアクセス時の流れ

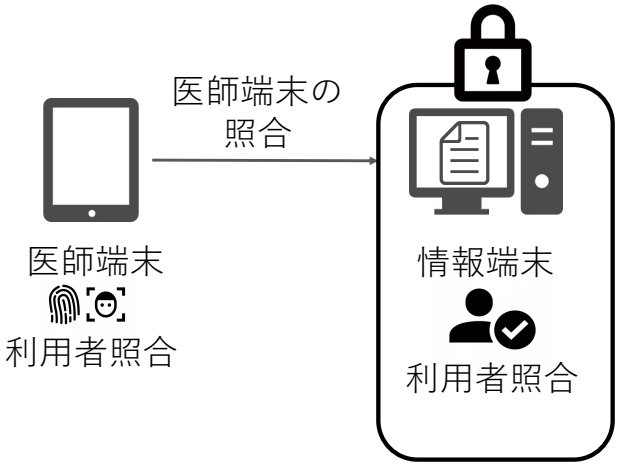

図 4.4　リモートアクセスに必要な照合

4.3.2　情報端末

情報端末はリファレンスモニタへ検索要求できる端末を制限するために安全が確保できる場所に設置する．リモートアクセスされた情報端末からリファレンスモニタへ検索要求を送信する流れを図 4.5 に示す．情報端末にリモートアクセスしてきた端末が正当な医師端末か

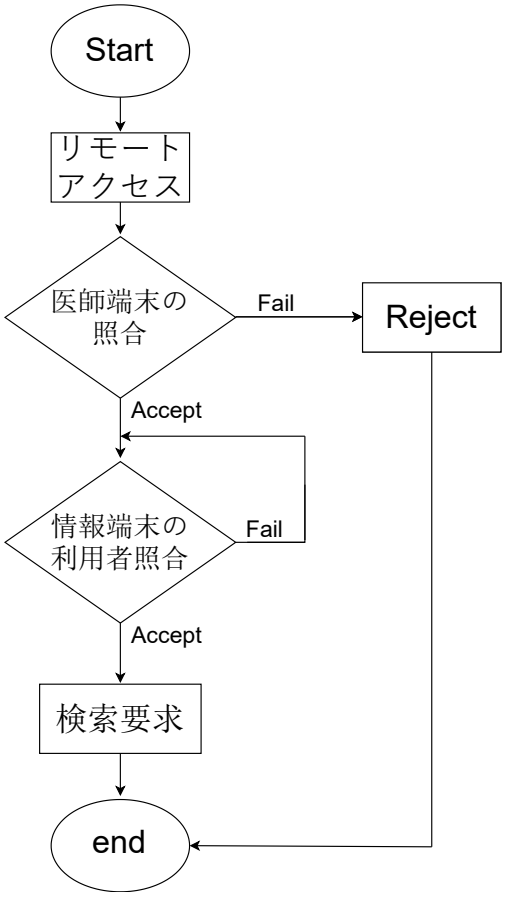

図 4.5　検索要求時の流れ

確認する必要がある．情報端末には事前に正当な医師端末情報を登録しておき，医師端末の照合を行う．また，情報端末にリモートアクセスしてきた利用者が正当な利用者であるか判断するために利用者照合する必要がある．利用者が正当であった場合，リファレンスモニタに検索要求を送信する．検索要求を行なうために必要な照合の様子を図 4.6 に示す．

4.3.3　リファレンスモニタ

ストレージへのアクセスを制限するためにストレージにアクセスできる端末をリファレンスモニタとし，安全が確保できる場所に設置する．また，リファレンスモニタが検索・復元を行う．そのため，検索・復元に必要な $p(tag), x_u(tag)$ が保存される．ストレージ群へのアクセス時の流れを図 4.7 に示す．リファレンスモニタにアクセスしてきた端末が正当であ

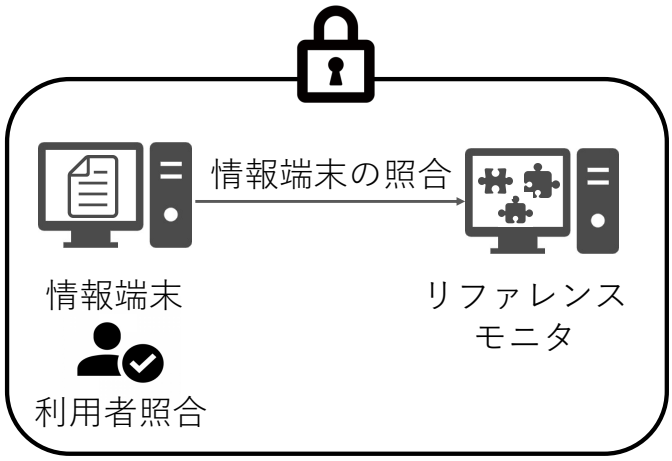

図 4.6　検索要求に必要な照合

るか判断する必要がある．事前にリファレンスモニタに正当な情報端末を登録しておき，情報端末の照合を行う．リファレンスモニタの利用者照合は，情報端末での利用者照合がクリアされた場合，正当な利用者としてリファレンスモニタの利用が可能となる．そして，ストレージ群へアクセスする．ストレージ群へのアクセスに必要な照合の様子を図 4.8 に示す．

4.3.4　ストレージ群

ストレージ群は医療データのシェア，$x_u(med), t_u$ を保存しているため，アクセスできる端末を制限する必要がある．そのため，アクセスできる端末をリファレンスモニタに制限し，アクセスしてきた端末の照合を行う．事前に正当なリファレンスモニタを登録しておき，リファレンスモニタの照合を行う．また，ストレージ群にアクセスしてきた利用者が正当であるか照合を行う．照合がクリアされた場合，ストレージ群に保存されているデータをリファレンスモニタに提供する．リファレンスモニタへ医療データ提供時の流れを図 4.9 に示す．

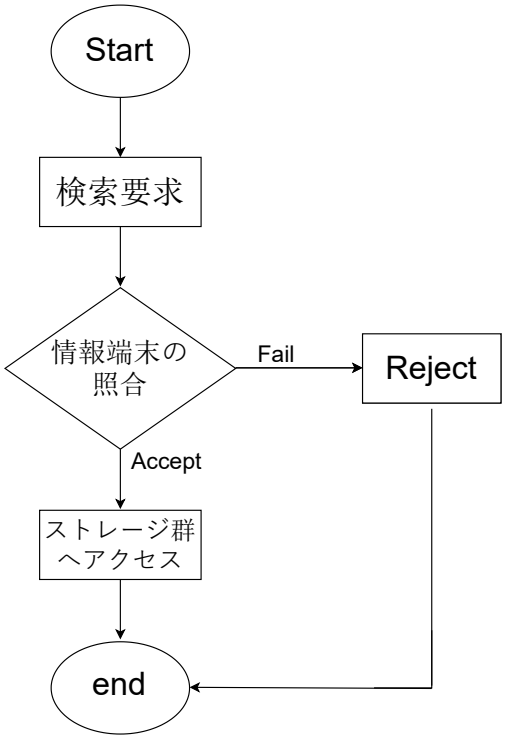

図 4.7　ストレージ群へのアクセス時の流れ

4.4　評価

検索システム構成の評価として，検索システム構成に対して想定されるリスクについて述べる．

検索システム構成は，医師端末の安全が保証できないことから医師端末が盗難されることや医師の不注意等によってその他の人物による医師端末利用が考えられる．また，医師端末，情報端末やリファレンスモニタのなりすましが考えられる．

医師端末が盗難された場合

検索・復元を行う端末をリファレンスモニタとし，$p(tag), x_u(tag)$ をリファレンスモニタに保存しているため，医師端末の盗難による $p(tag), x_u(tag)$ の漏えいを防止できる．また，医師の不注意等により，その他の人物によって医師端末の利用が可能な場合や医師端末のなりすましが行われた場合，情報端末に対してリモートアクセスの実行が試みられる．しかし，情報端末では，医師端末の利用者が正当な人物であるか確かめるため

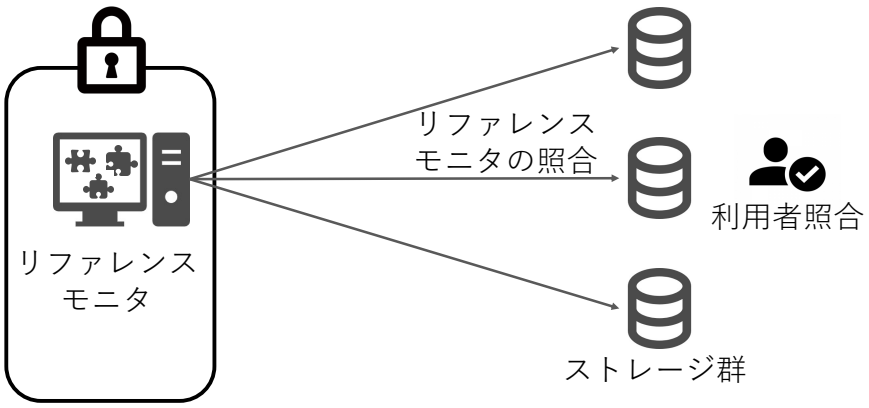

図 4.8　ストレージ群へのアクセスに必要な照合

の利用者照合があるため情報端末にリモートアクセスすることはできない．

リファレンスモニタへ検索要求が送信された場合

情報端末のなりすましによってリファレンスモニタへ検索要求が送信された場合，情報端末での利用者照合がクリアされていない場合，リファレンスモニタにおいて利用者照合が行われるため，不正な検索要求による医療データの取得はできない．

ストレージ群へアクセスされた場合

リファレンスモニタのなりすましによってストレージ群へアクセスが行われた場合，医療データのシェアと対応する t_u は漏えいする．しかし，医療データのシェアの組み合わせがわからないようにストレージに保存する方法を用いている場合，医療データのシェアの組み合わせを知ることができない．また，医療データのシェアの対応を知る t_u からも医療データのシェアの組み合わせを知ることはできないため，医療データを不正復元される恐れはない．

検索システム構成に対して評価を行った図を図 4.10 に示す．検索システム構成により安全性を担保できると考える．

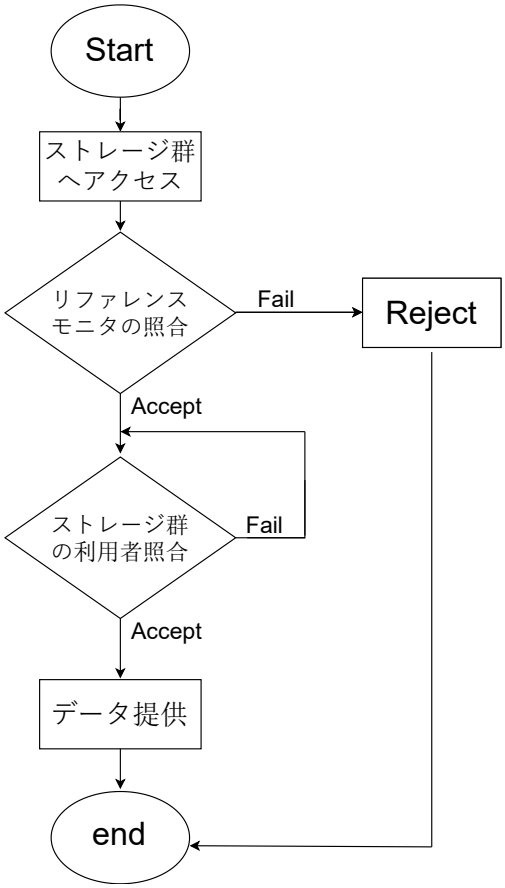

図 4.9　リファレンスモニタへ医療データ提供時の流れ

4.5　まとめ

検索方法により識別情報が絞られる恐れがある．剰余演算の法を大きくすることや医療データを不正復元する手順より識別情報の絞り込みは大した問題ではない．しかし，識別情報の絞り込みによって医療データが不正復元されるきっかけになるようなことがあってはならない．本章では，識別情報の絞り込みの要因となる $t_u, p(tag), x_u(tag)$ を検索システム構成により漏えいを防止するための検索システムの構成要素について述べ，検索システム構成について述べた．検索システム構成では，ストレージ群からの t_u の漏えいを防止するため，ストレージ群へアクセスできる端末をリファレンスモニタに制限した．また，医師端末の盗難による $p(tag), x_u(tag)$ の漏えいを防止するため，検索・復元を行う端末をリファレンスモニタとし，$p(tag), x_u(tag)$ を保存し，リファレンスモニタへのアクセスを情報端末に制限

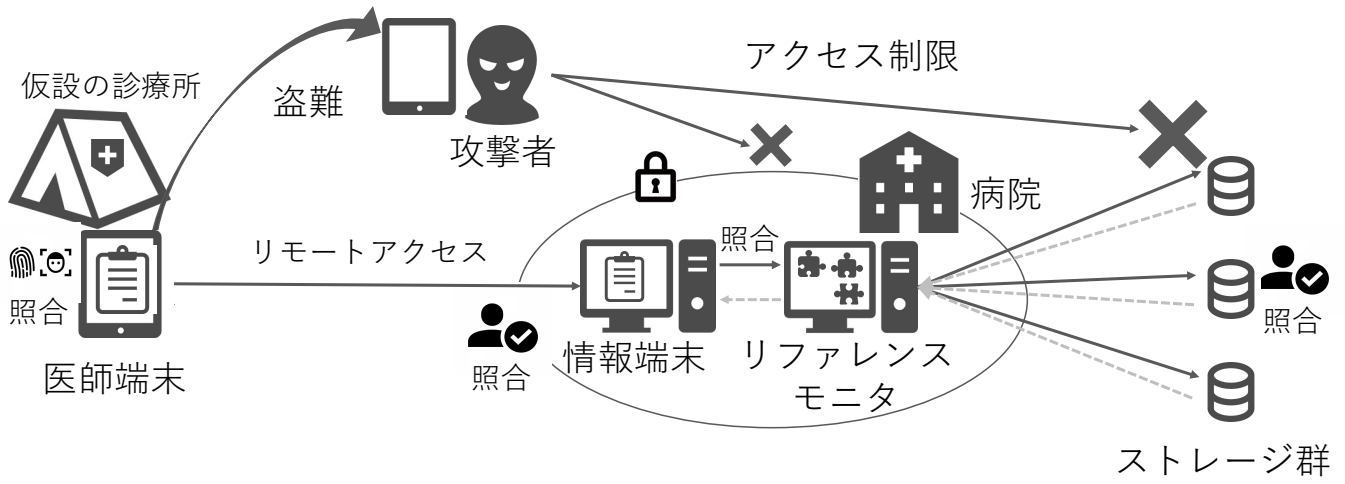

図 4.10　検索システム構成の評価

した．検索システム構成により安全性を担保できると考える．

第 5 章

結び

5.1 本研究のまとめ

本研究では，検索方法によって生じた識別情報の絞り込みの防止を目的として検索システム構成の提案を行い，評価した．

検索システム構成では，ストレージ群へのアクセスによる t_u の漏えいを防止するため，ストレージ群へアクセスできる端末をリファレンスモニタに制限した．また，リファレンスモニタのなりすましによってストレージ群にアクセスされても利用者照合や医療データのシェアと t_u の保存方法により，医療データのシェアの組み合わせを知ることはできない．医師端末の盗難による $p(tag), x_u(tag)$ の漏えいを防止するため，検索・復元を行う端末をリファレンスモニタとし，$p(tag), x_u(tag)$ を保存した．リファレンスモニタへの検索要求を情報端末に制限した．また，情報端末のなりすましによってリファレンスモニタに検索要求が送信されても情報端末での利用者照合がクリアされていない場合，リファレンスモニタで利用者照合が行われるため，不正な検索要求による医療データの取得はできない．

検索システム構成により，安全に医療データの検索方法を使用することができる．

5.2 今後の課題

本研究では，リモートアクセスや端末間の通信路の安全性については考えていないため，端末間の通信路についても考える必要がある．

謝辞

本研究を行うにあたり，終始ご指導頂きました高知工科大学情報学群の福本昌弘教授に謹んで感謝致します．理解力の足りない私に対して，最後まで見捨てず，ご指導頂き大変感謝しています．本研究の副査をして頂いた情報学群敷田幹文教授，原田崇司助教のお二人にも謹んで感謝致します．原田先生には，何度もお食事に誘って頂いたり，勉強会で大変お世話になりました．

NOC の職員であり，研究室の OB でもある福冨英次氏にも謹んで感謝致します．何度もお食事や遊びに誘って頂きました．これからもシェフの育成をかんばってください．

修士 2 年の中村巴氏にも謹んで感謝致します．研究についてわからないことを何度も丁寧に教えて頂きました．また，私の読みづらい文章を何度も添削 (赤ペン先生) して頂き，ありがとうございました．研究を最後までやり遂げることができたのは中村巴氏のおかげです．今後は横浜のオシャレなお店でお食事しましょう．

修士 1 年の小野田祐稀氏にも謹んで感謝致します．研究室イベントでは，何度もお手伝いして頂きました．これからも頑張って下さい．

同期の松本侑馬氏，斉藤直弥氏にも感謝しています．2 人がいなければ，卒業研究を乗り越えることはできなかったと思います．2 人のおかげで楽しい研究室生活でした．

福本研究室 3 年生のみなさん，今年はコロナウィルスの影響で一緒に活動できる時間が少なく残念でした．これから就活や卒業研究など大変なことがあると思いますが，無理せず，頑張って乗り越えて下さい．

最後になりましたが，4 年間の学生生活を支えてくださった家族や友人などの皆様に感謝致します．また，高知で出会った全ての人に感謝致します．

参考文献

[1] 福本昌弘，“高知県における電子カルテ遠隔バックアップと部分復元可能な秘密分散法，”北隆館，Precision Medicine，pp.57-61，Vol.3，No.9，2020.

[2] 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，pp.31-36，Dec.2015.

[3] 中村巴，福冨英次，福本昌弘，“部分復元可能な秘密分散法におけるシェア間の係数の差の比較による医療データの検索，”電子情報通信学会 暗号と情報セキュリティシンポジウム (SCIS2020)，3C1-4，2020.

[4] 厚生労働省，“医療情報システムを安全に管理するために，”https://www.mhlw.go.jp/shingi/2009/03/dl/s0301-6b.pdf，2021 年 2 月 10 日閲覧.

[5] 木村通男，“HL7 入門 SS-MIX ストレージ，”http://www.hl7.jp/docs/55seminar\verb|_|1\verb|_|HL7.pdf，2021 年 2 月 10 日閲覧.