平成 26 年度

学士学位論文

秘密分散データの部分秘匿化復元の検証

Verification of partly concealed reconstructing of secret sharing data

1150324　　田中 麻実

指導教員　　福本昌弘

2015 年 3 月 4 日

高知工科大学 情報学群

要 旨

秘密分散データの部分秘匿化復元の検証

田中 麻実

秘密分散法を用いて個人情報が含まれる医療データを分散バックアップするという取り組みがある．平常時は医療機関ごとに医療データのバックアップを行い，災害時にバックアップデータを活用しようとするものである．医療データには個人情報が含まれているため，災害時派遣された医療従事者が診察に不必要なデータを閲覧することは問題である．バックアップデータから患者の診察に必要なデータのみを閲覧することができれば，外部から派遣された医師は患者に対して適切な処置を行うことができる．しかし，秘密分散法はデータを部分的に秘匿したまま復元することはできない．

そこで本研究では，(k, n) しきい値秘密分散法を用いてデータを部分的に秘匿したまま復元する方法を提案し，提案方法の評価を行っている．データを分割してからシェアを作成し，シェアを結合する際に秘匿部分にマスクをかけることで，データを部分的に秘匿している．しかし，何らかの方法でシェアを入手することができれば，データを不正に復元される恐れがあるため，不正復元を防ぐアクセス制限方法を示している．アクセス制限により，管理者の意図しないユーザがデータを復元することを防ぐことができる．

キーワード　秘密分散法，部分秘匿化復元

Abstract

Verification of partly concealed reconstructing of secret sharing data

TANAKA Asami

To prevent medical data from loss and leak, there is the distributed backup system using secret sharing scheme. In this system, medical data ware backed up to network storages for each medical institution in normal time. When the disaster occurs, backup data is used for medical examination to victims. Disaster Medical Assistance Team (DMAT) should not be see data other than needed data in the medical examination. DMAT can examine and treat when needed data can be seen. However, secret sharing scheme cannot conceal a part of data when distributed data is reconstructed.

In this paper, the reconstruction method to conceal a part of data using (k, n)-threshold secret sharing scheme has been proposed and evaluated. Data is divided private part and public part. In proposed method, shares are made from divided data. Shares in private part are concealed by dummy data when shares are bonded, and bonded shares are reconstructed. However, shares are obtained fraudulently, data can be reconstructed. As a solution against this vulnerability, the access limitation method of shares have been represented. The access limitation can be prevented from illegal reconstructed.

key words secret sharing scheme, partly concealed reconstructing

目次

図目次

表目次

第 1 章

序論

本章では, 本研究の背景と目的を述べ, その後本論文の構成について述べる.

1.1 本研究の背景と目的

広域災害に備えて，電子カルテなどの医療データをネットワーク上のストレージにバックアップする取り組みがある．その中の一つにデータの秘匿化，冗長化ができる秘密分散法を用いて，個人情報が含まれる医療データをネットワーク上の複数のストレージに分散バックアップするという取り組みがある．これは平常時に医療機関ごとに保持している医療データを秘密分散法を用いてネットワーク上の複数のストレージに分散バックアップを行い，災害時にバックアップしたデータを活用しようとする取り組みである．

災害時は負傷者が多い上にその地域の医療機関も被災している可能性が高い．そのため，その地域の医療機関や，医療機関に所属している医師や看護師などの医療従事者だけでは，負傷者の対応することは困難である．そこで，災害発生直後に迅速に動けるような訓練を受けた DMAT(Disaster Medical Assistance Team) と呼ばれる災害派遣医療チームが手助けに入り，DMAT に続いて日本医師会が結成した JMAT(Japan Medical Assistance Team) などの医療チームが手助けに入る．

DMAT や JMAT などの外部から派遣された医療従事者は，負傷者や被災者のプライマリケアを行う．その際に，バックアップした医療データが存在すれば，外部から派遣された医療従事者は患者に対して適切な処置を行うために医療データ閲覧の要求を行う．しかし，医療データは個人情報が含まれているため，患者本人が同意していない限り第三者に閲覧させ

ることはできない．患者の診察に必要なデータのみをバックアップデータから抜き出すことができれば医療データの開示を行うことができるが，秘密分散法で分散バックアップしたデータの復元は，復元できるかできないかのどちらかであり，患者の診察に不必要なデータを秘匿して復元することはできない．そこで，秘密分散法を用いて分散バックアップしたデータを部分的に秘匿したまま復元する方法を考える必要がある．

本研究では，(k, n) しきい値秘密分散法で分散されたデータを部分的に秘匿したまま復元する方法を提案し，提案方法の評価を行う．

1.2　本論文の構成

本節では本論文の構成を述べる．2 章では，医療データの分散バックアップについて述べる．バックアップ対象の医療データについて述べ，分散バックアップに用いる秘密分散法を選定する．

3 章では，部分秘匿化復元について述べる．まず初めに，バックアップした医療データの活用方法について述べる．次に，部分秘匿化復元について述べる．最後に，シェアに対してのアクセス制限を述べる．

4 章では，部分秘匿化復元とシェアに対するアクセス制限の評価を行う．それぞれの評価を行い，最後に考察を行う．

5 章では，本研究をまとめ，今後の課題を述べる．

第 2 章

医療データ分散バックアップ

東日本大震災で岩手県の沿岸部にある医療機関の医療データが津波によって喪失した．医療データが喪失したことにより，患者がどのような病気でどのような薬剤が処方されていたかが分からず，適切な処置を行うことが困難になるという問題が起きた．これはプライマリケア復旧の大きな妨げになった [1].

高知県がうける南海トラフ大震災による津波被害は甚大と予想されている．また，高知県の主要な病院は人口が密集している海沿いにあるため，岩手県のように医療データが津波によって喪失する可能性が高い．そこで，高知県では医療データを県外にバックアップする計画が進められている．図 2.1 のように高知県の県民の 6 割が利用している主要な病院 13 ヶ所が参加し，高知県へき地医療情報ネットワークを構成し，JGN-X(Japan Gigabit Network eXtreme) を用いて県外へバックアップを行う計画である．現在は，高知県全体の医療データを一ヶ所に集め，県外一ヶ所にバックアップをとっている [2].

一ヶ所にバックアップすることで，災害時にバックアップデータを保存しているストレージが故障や破損していた場合，データをリストアすることができない．バックアップデータを一ヶ所のストレージに保存するのではなく，複数のストレージに保存することで，バックアップデータを冗長化することができる．しかし，冗長化を行うと情報漏えいのリスクが上がる．そこで，データを分散バックアップする方法として，秘密分散法がある [3]．過去に秘密分散法を用いてファイルシステムを実装し，分散バックアップに秘密分散法が適していると示している研究がある．[4]．また実際に，医療データを分散バックアップ手段として提案されている [5].

本章では医療データの分散バックアップに用いる秘密分散法の選定を行う．まず，バック

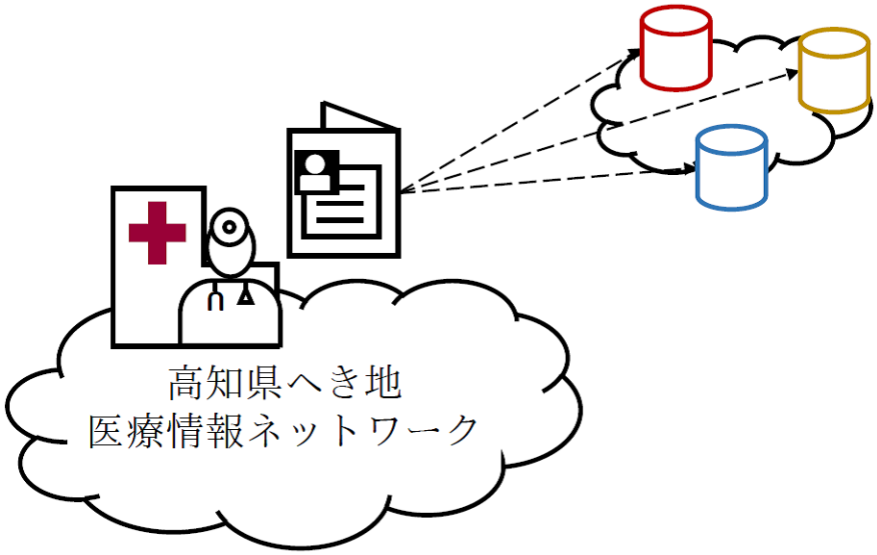

図 2.1　秘密分散法を用いた医療データバックアップ

アップ対象の医療データについて述べ，医療データを分散バックアップするのに必要な条件を述べる．次に，秘密分散法で代表的な，(k, n) しきい値秘密分散法を述べ，(k, n) しきい値秘密分散法を拡張した (k, d, n) ランプ型しきい値秘密分散法について述べ，分散バックアップの選定基準とする秘匿関数計算とアクセス制限について述べる．最後に (k, n) しきい値秘密分散法と (k, d, n) ランプ型しきい値秘密分散法を比較し，医療データの分散バックアップに適している秘密分散法を決定する．

2.1　バックアップ対象の医療データ

本節では，バックアップ対象の医療データについて述べる．医師が患者を診察するのに必要不可欠なカルテを電子化した電子カルテと，医療機関が保険者に対して月単位で発行している医療報酬の明細表であるレセプトについて述べる．

電子カルテ

電子カルテとは，一般に医療機関で医療従事者が記録するカルテをコンピュータなどを用いて電子的に記録保存するシステムのことを指す．

2.1 バックアップ対象の医療データ

紙のカルテを利用するのに比べ，保存や保管が容易でありネットワークさえつながっていれば，病院内のあらゆる場所で読み出しや，書き込みができる．カルテは図や話し言葉，記号など様々な形式で書かれることが多く，患者の名前や年齢のように簡単に構造化して電子カルテに組み込むことはできないため，紙のカルテほど自由に書き込むことはできない．また，電子媒体であるため，停電時には使用することができない．

ネットワークさえつながっていれば，電子カルテにアクセスすることができるため，改ざんやなりすまし，情報漏えいなどの第三者からの攻撃に対するセキュリティ面の対策が必要になる．

カルテは，医療法，歯科医師法に基づいて作成される公的な書類であり，紙媒体による管理が義務付けられていたが，1999 年に厚生省がカルテの保存を緩和させたことにより電子保存が可能になった．また，厚生省は電子保存の三原則も定めた．電子保存の三原則は，電子カルテの三原則ともよばれ，真正性，見読性，保存性から構成されている [6].

真正性

真正性とは，医師が記録，または確認したデータに関して，第三者から見て，作成の責任の所在が明確であり，故意または過失による虚偽の入力や書き換え，消去などが防止されていることである．

見読性

見読性とは，電子媒体に保存した医療データが診察に用いることに支障がないことや監査などに差し支えがない内容である上に，患者などの権限保有者からの要求に基づき，必要に応じて見読可能な状態にすることである．

保存性

保存性とは，電子カルテの情報が法律や法令などに定められた期間の間真正性を保ち，見読可能である状態で保存されていることである．

電子カルテを使用する際はこの三原則を守らなければならない．

医療データでもっとも重要な電子カルテのフォーマットは各ベンダーで作成されてい

るが，互換性がなく，標準規格も定まっていない．厚生労働省は，厚生労働省電子的診療情報高官推進事業 (SS-MIX) を用いて医療データの標準化を目指しているが，導入にコストがかかるため，医療機関ではあまり普及していない．また，電子カルテの標準規格が定まっていないため，データの読み込みにはその規格に対応した機材が必要になる．そのため，災害時病院の機能が復旧してからでないと電子カルテを読むことは難しい．

レセプト

レセプトは，医療機関が保険者に月単位で発行する医療報酬の明細表である．標準規格が定まっていない電子カルテとは異なり，データ形式がテキストデータに統一されている．レセプトの内容は，発行する医療機関によって多少異なるが，患者の氏名などの個人情報や，患者の健康保険加入情報，請求元の医療機関名やその医療機関での診察や処方された薬剤の情報が載っている．レセプトは紙媒体で発行されていたが，電子レセプト化が進み 2011 年には電子レセプトの普及率が 94 パーセントになっている．そのため，おおよその医療機関では電子レセプトを活用していると言える [7]．電子レセプトが普及しているため，各医療機関で蓄積されたレセプトを分析活用することに期待が高まっているが，レセプトは保険請求に使用されていることを目的として作成されているので，簡単に分析できる構造になっていない．

東日本大震災の際に，DMAT として被災地に赴いた医師は，電子カルテが参照できなくとも直近の診療記録があれば的確な医療行為が可能であると述べている [1]．レセプトにはその月の診療内容が記載されているため，災害発生時に医師がレセプトを活用することによって患者に対して適切な処置を行うことができるのではないかと期待されている．

電子カルテは，患者を診察する上で必要不可欠なものだが，標準規格が定まっておらず，規格ごとに読めるソフトウェアが入った機材が必要になるため，災害発生直後に使用することは難しい．そこで，医療機関が保険者に対して月単位で発行しているレセプトが注目され

ている．レセプトは，テキストデータに統一されているため，特別なソフトウェアや機材は必要ない．そのため，災害発生時にレセプトを活用して患者に対して適切な処置を行えるのではないかと注目されている．電子カルテ，レセプトともに患者の個人情報を含んでいるため，アクセス制限をかけるなど取扱いは注意しなければならない．

2.2　医療データを分散バックアップするための条件

前節では，バックアップする医療データについて述べた．本節では，医療データを分散バックアップするための条件を述べる．

まず，広域災害に備えてバックアップを行うため，バックアップしているネットワーク上のストレージが一台故障しても復元できるようにデータを冗長化する必要がある．また，医療データは個人情報を含むため，分散バックアップデータから情報を読み取れないように秘匿化しなければならない．次に，元の医療データにアクセスする権限がない人がバックアップデータにアクセスできると，バックアップデータをリストアすることにより元の医療データにアクセスすることができる．よってバックアップデータにも元の医療データと同様のアクセス制限をかけなければならない．最後に，バックアップデータを安全に利活用するために，秘匿化されたバックアップデータから医療データの内容をデータ解析者に知られることなくデータ解析が可能な秘匿関数計算ができればよい．

医療データを分散バックアップするための条件をまとめる．

1. バックアップデータの冗長化ができること
2. バックアップデータの秘匿化ができること
3. バックアップデータにアクセスできる人を制限できること
4. 秘匿されたバックアップデータから医療データのデータ解析ができること

以上の条件を満たす分散バックアップ方法を選定する．

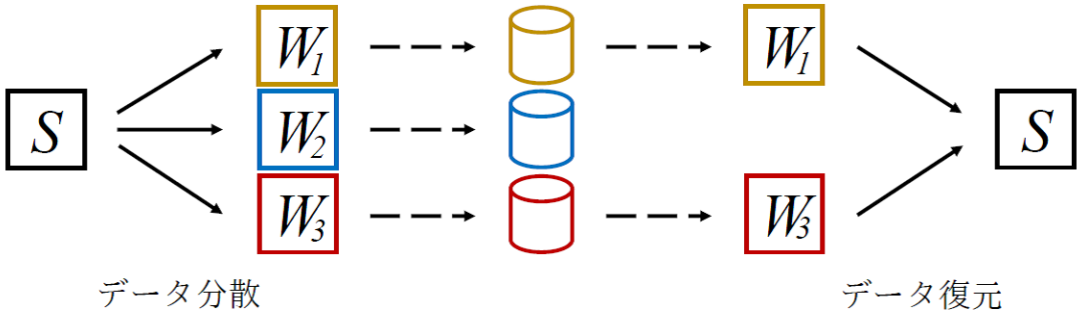

図 2.2 (k, n) しきい値秘密分散のデータの流れ

2.3 (k, n) しきい値秘密分散法

本節では，分散バックアップに用いられる秘密分散法の一つである (k, n) しきい値秘密分散法について説明を行い，医療データの分散バックアップに適しているかどうかを検討する．

(k, n) しきい値秘密分散法は，図 2.2 のようにデータを n 個の分散データ (以下シェアと呼ぶ) にし，シェアを k 個以上集めることでデータを復元できる方法である [3]．n 個のシェアのうち k 個以上を集めることでデータの復元が可能になり，k 個未満であればデータの情報が得ることができない性質を持つ．これは，情報理論的に示されている [8]．(k, n) しきい値秘密分散法を用いて作成したシェアは，単一のシェアからデータの情報を得ることができないためデータ全体の秘匿化が可能である．また，n 個のシェアのうち $n - k$ 個紛失したとしても復元できるためデータの冗長化ができる．シェアのデータサイズは，多項式補完に基づいて作成されているため，元データのデータサイズと等しい．よって，n 個のシェア全体のデータサイズは元のデータのデータサイズの n 倍となり，符号効率が悪いと言える [9]．

(k, n) しきい値秘密分散法を用いたデータの分散方法，復元方法を以下に示す．

データを S，素数を p とし，n 個のシェアを $w_i (i = 1, \cdots, n)$ を作成する．

分散方法

1. データの管理者は $S < p$ かつ $n < p$ である任意の素数 p を選ぶ．
2. $\mathbb{Z}/p\mathbb{Z}$ から異なる n 個の x_i と $k - 1$ 個の乱数 $r_j (j = 1, \cdots, k - 1)$ を選択する．

3. シェア w_i を以下の式 2.1 にて作成する.

$$w_i = S + r_1 x_i + \cdots + r_{k-1} x_i^{k-1} \pmod{p} \tag{2.1}$$

4. x_i をシェアの ID として, x_iと w_i をセットにする.

5. x_i と w_i のセットを n 個のストレージに保管する.

復元方法

1. n 個のストレージから w_i と ID のセットを k 個以上取得する.

2. 式 2.1 に代入し, k 個の線形方程式を求める.

$$\begin{aligned} S + r_1 x_1 + r_2 x_1^2 + \cdots + r_{k-1} x_1^{k-1} &= w_1 \pmod{p} \\ S + r_1 x_2 + r_2 x_2^2 + \cdots + r_{k-1} x_2^{k-1} &= w_2 \pmod{p} \\ &\vdots \\ S + r_1 x_n + r_2 x_n^2 + \cdots + r_{k-1} x_n^{k-1} &= w_n \pmod{p} \end{aligned} \tag{2.2}$$

3. k 個の方程式から S と r_j を求めることで, 元データを得ることができる.

式 2.2 は方程式であるため, 行列で表すことができる.

$$\begin{pmatrix} 1 & x_1 & x_2^2 & \cdots & x_1^{k-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_2^2 & \cdots & x_n^{k-1} \end{pmatrix} \begin{pmatrix} S \\ r_1 \\ \vdots \\ r_{k-1} \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \pmod{p} \tag{2.3}$$

(k, n) しきい値秘密分散法は，作成した n 個のシェアを $n - k$ 個紛失したとしてもデータを復元することができるため，データの冗長化がとれる．また，作成したシェアから元データの情報を読み取ることができない上に，シェアを k 個未満集めたとしても元データの情報を得ることができないため，データの秘匿化が可能である.

2.4 (k, d, n) ランプ型しきい値秘密分散法

(k, n) しきい値秘密分散法で作成したシェア n 個全部のシェアのデータサイズは，元データのデータサイズの n 倍になるため符号効率が悪い．そこでシェアの符号効率を改善する

ために考案されたのが，(k, d, n) ランプ型しきい値秘密分散法である．本節では，(k, d, n) ランプ型しきい値秘密分散法が，データの分散バックアップに適しているかどうかを検討する．

(k, d, n) ランプ型しきい値秘密分散法は，n 個のシェアのうち任意の k 個が得られればデータを復元でき，任意の $k-d$ 個では全くデータの情報が得られず，任意の $k-t(1 \leq t \leq d-1)$ 個のシェアでは，t が小さくなるにつれて段階的にデータの一部の情報が得られるようになる性質をもつ秘密分散法である [10]．(k, d, n) ランプ型しきい値秘密分散法は，(k, n) しきい値秘密分散法の符号効率を改善したものである．(k, n) しきい値秘密分散法で作成したシェアのデータサイズは元データと同じデータサイズなのに比べ，(k, d, n) ランプ型しきい値秘密分散法で作成したシェアのデータサイズは元のデータサイズよりも $\frac{1}{d}$ となっている．

(k, d, n) ランプ型しきい値秘密分散法の分散方法を以下に示す．

データ S を d 個に分割する．d は $d < k$ となるように定める．

$$S = S_1 + \cdots + S_{d-1} \tag{2.4}$$

次に，データの部分的な値 $S_l(l = 0, \cdots, d)$ と素数 p，乱数 $r_j(j = 1, \cdots, k-d)$ を用いて式 2.5 でシェアを求める．

$$w_i = S_0 + S_1 x + \cdots + S_{d-1} x^{d-1} + r_1 x^d + \cdots + r_{k-d} x^{k-1} \pmod{p} \tag{2.5}$$

復元時には，データを復元できるようにシェアを k 個以上集め，式 2.5 に当てはめると，復元することができる．

(k, d, n) ランプ型しきい値秘密分散法は，(k, n) しきい値秘密分散法と同じように，データを冗長化，秘匿化することができる．(k, n) しきい値秘密分散法で作成したシェアよりも $\frac{1}{d}$ だけ符号効率が良くなったが，$k-t$ 個シェアを集めた場合，データの一部分の情報が得られる．

2.5　秘匿関数計算

秘匿化を行ったデータを解析する際に，秘匿化を解除してデータ解析を行うと，データ解析者にデータの内容を知られてしまう．データ解析者にデータの内容を知られることなくデータを解析する技術として秘匿関数計算技術がある．安全に医療データを解析するために秘匿関数計算は必要である．秘密分散法を用いて作成したシェアが秘匿関数計算を行うことができるかを医療データの分散バックアップに用いる秘密分散法の決定の一基準とする．本節では，元データの値を知ることなく，秘密分散法で作成したシェアから元データ同士の演算や，元データのデータ処理などを行うことができる秘匿関数計算について述べる．

秘匿関数計算とは，データの値を知ることなく演算やデータ処理などを行う技術のことである [11]．例えば，(k, n) しきい値秘密分散法において，データ S_1, S_2 のシェアをユーザ A が持っていたとする．ユーザ A が所持するシェアを $w_{1,j}, w_{2,j}$ とすると，$w_{1,j} + w_{2,j} \pmod p$ は $S_1 + S_2 \pmod p$ のシェアになっている．シェアを作成する多項式 2.1 にあてはめると

$$w_1 = s_1 + r_{1,1}x_i + \cdots + r_{1,k-1}x_i^{k-1} \pmod p \tag{2.6}$$

$$w_2 = s_2 + r_{2,1}x_i + \cdots + r_{2,k-1}x_i^{k-1} \pmod p \tag{2.7}$$

となるから，$w_1 + w_2$ の和は

$$w_1 + w_2 = s_1 + s_2 + x_i(r_{1,1} + r_{2,1}) + \cdots + x_i^{k-1}(r_{1,k-1} + r_{2,k-1}) \pmod p \tag{2.8}$$

となる．$w_1 + w_2$ を復元すると $S_1 + S_2$ を得ることができる．よってユーザ A はデータを復元することなく $S_1 + S_2$ を得ることができる．準同型をもつ準同型暗号や (k, n) しきい値秘密分散法を用いて作成したシェアは秘匿関数計算を行うことができる．

2.6　アクセス制限

前節で述べた医療データは個人情報が含まれるデータであるため、閲覧や書き込みなどアクセスできる人を制限しなければならない．本節では，ユーザやプロセスがデータにアクセス可能か不可能かを判断するアクセス制限について述べる．

2.6　アクセス制限

アクセス制限とは，あるサブジェクトがどのオブジェクトに対してアクセス可能か不可能かを判断する機能である．

サブジェクト

アクセスを主体的に行うものであり，主体とも呼ばれる．また複数のプロセスをまとめてサブジェクトと呼ぶこともある．サブジェクトは一般的にユーザやプロセスを指す．

オブジェクト

アクセスを受けるものであり，対象とも呼ばれる．オブジェクトは一般的にファイルを指す．

アクセス

サブジェクトがオブジェクトに対して何らかの操作を行うことである．主な操作として，参照，更新，生成，削除などが挙げれる．

アクセス許可

サブジェクトがオブジェクトに対して許可されている操作のことである．アクセス権限とも呼ばれることもある．

アクセス制御ポリシー

どのサブジェクトがどのオブジェクトにアクセスするかの規則の集合である．アクセス制御ポリシーには様々な種類がある．ここでは重要度が厳密につけられたオブジェクトによく用いられる階層的アクセス制御ポリシーと，重要度が並列したオブジェクトに用いる多元的アクセス制御ポリシーについて述べる．

階層的アクセス制御ポリシー

階層的アクセス制御ポリシーとは，アクセス制御ポリシーの一つであり，データを階層付け，階層間のアクセスを制限するものである．この制御ポリシーは，データに異なる重要度が設定されているような場合に用いられることが多い．縦方向にデータのアクセスを制限するため，重要度が厳密につけられた情報のアクセス制限には向いている．しかし，重要度が並列したデータのアクセス制限には向いていな

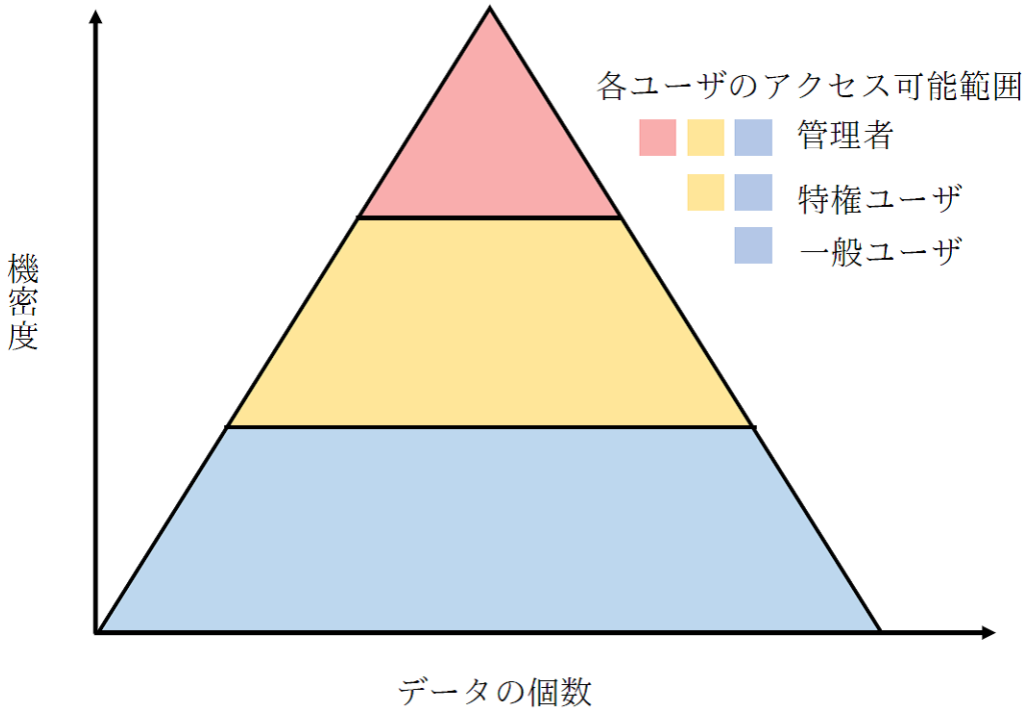

図 2.3　階層的アクセス構造

い．階層的アクセス構造を図 2.3 に示す．縦軸はデータの機密度を表しており，横軸はデータの数を示している．一般ユーザは青色に属するオブジェクトしかアクセスできず，特権ユーザは青色と黄色に属するオブジェクトにアクセスすることができる．管理者は全てのオブジェクトにアクセスできる．このように，データの機密度によってアクセス制限を行うことができる．

階層的アクセスポリシーモデルとして，BLP(Bell LaPadula) モデル，Biba Integrity モデルなどがある [12].

多元的アクセス制御ポリシー

多次元的アクセス制御ポリシーとは，アクセス制御ポリシーの一つであり，データを区切り，区切った情報へのアクセスを制限する方法である．階層的アクセス制御ポリシーでは縦方向にデータアクセスを制限をしていたが，多元的アクセス制御ポリシーは横方向にデータアクセスを制限する．よって階層的アクセス制御ポリシーでは困難であった重要度が並列したデータのアクセス制限ができる．しかし，重要

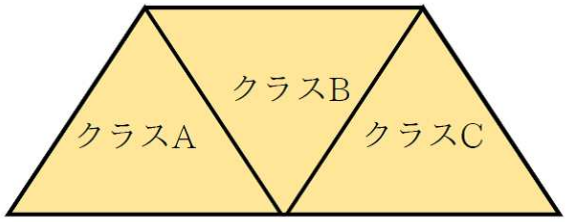

図 2.4　多元的アクセス構造

度が厳密につけられた情報のアクセス制御には向いていない．多元的アクセス構造を図 2.4 に示す．クラス A，クラス B，クラス C はすべて同じ重要度である．一つのクラスにアクセスできるサブジェクトは他のクラスにアクセスすることができないというようなアクセス制限を行うことができる．

多次元的アクセスポリシーモデルとして，Clark Wilson モデル，Chinese-Wall モデルなどがある [12].

リファレンスモニタ

アクセス制御ポリシーを用いて，サブジェクトからのアクセス要求に対してアクセス可能かアクセス不可能かを判定するものである．リファレンスモニタが正しく的確なアクセス制限を行うために，全てのアクセス要求に対して必ず呼び出されること，メカニズムは破壊や改ざんを受けないこと，正しく動作することが保証されていることの 3 つがリファレンスモニタ実装の要件となっている [12].

アクセス制限がかかっているオブジェクトに対するアクセス手順の流れを図 2.5 に示す．アクセス手順は以下の通りである．

1. サブジェクトはリファレンスモニタに対してアクセス要求を送る
2. 要求を受けたリファレンスモニタは，アクセス制御ポリシーが格納されているデータベースからアクセス制御ポリシーを読み出す
3. リファレンスモニタはアクセス制御ポリシーに基づき，サブジェクトがオブジェクトに対してアクセス可能か不可能かの判断を行う
4. アクセス不可能の場合はユーザからの要求を破棄して終了する．アクセス可能な場合

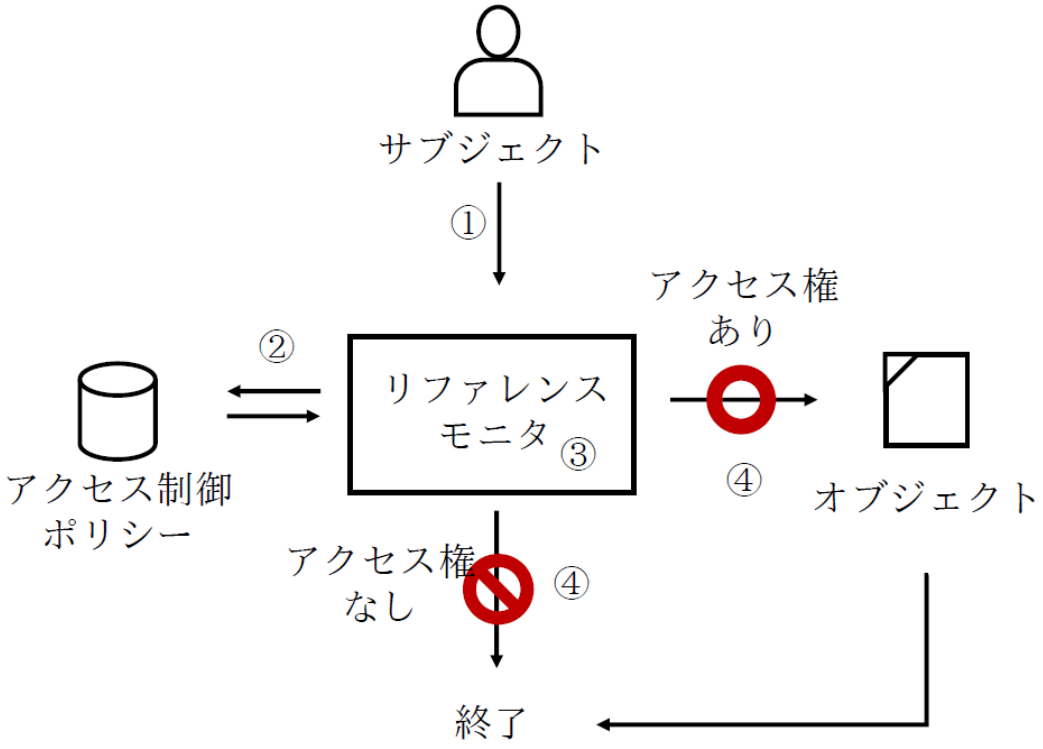

図 2.5　アクセス制限の流れ

は，サブジェクトはオブジェクトに対してアクセスの実行を行う

アクセス制限をかけることにより，オブジェクトにアクセスできるサブジェクトを制限することができる．

2.7　秘密分散法の比較

本節では，医療データの分散バックアップに適している秘密分散法を決定するために，(k, n) しきい値秘密分散法と (k, d, n) ランプ型しきい値秘密分散法の比較を行う．比較した結果を表 2.1 に示す．

(k, n) しきい値秘密分散法は，データを n 個のシェアにし，n 個のシェアのうち k 個以上集めることによってデータを復元できる．一つのシェアからは元データの情報を読み取ることはできず，シェアを k 個未満集めても元データの情報を得ることはできない．よって，

データの秘匿化ができる．また，n 個のシェアのうち $n-k$ 個紛失したとしてもデータを復元できるため，データの冗長化が可能である．(k,n) しきい値秘密分散法を用いて作成したシェアをあるアクセス制御ポリシーで割り当てアクセス構造を作ることができるため、シェアに対してのアクセス制限ができる [13]．(k,n) しきい値秘密分散法を用いて作成したシェアは，準同型を持つため，秘匿関数計算ができる．しかし，(k,n) しきい値秘密分散法は，多項式補完を用いてシェアを作成しているため符号効率が悪い．

一方，(k,d,n) ランプ型しきい値秘密分散法は，(k,n) しきい値秘密分散法で作成されたシェアの符号効率を改善したもので，(k,n) しきい値秘密分散法と同じようにデータの冗長化ができる．また，(k,d,n) ランプ型しきい値秘密分散法を用いて作成したシェアに対して，あるアクセス制御ポリシーを適応してアクセス構造を作成することができるため，シェアに対してアクセス制限をかけることができる．(k,d,n) ランプ型しきい値秘密分散法で作成したシェアのデータサイズは $\frac{1}{d}$ になり，(k,n) しきい値秘密分散法よりも符号効率が良い．しかし，シェアを $k-t(1 \leq t \leq d-1)$ 個集めた場合，元データの部分的な情報を得ることができる．よって，(k,n) しきい値秘密分散法よりもデータの秘匿性は低いといえる．また，(k,d,n) ランプ型しきい値秘密分散法は，必ずしも準同型を持たないため，秘匿関数計算を行うことができない場合がある．

よって，k 個未満集めても元データの情報が得られないかつ，作成したシェアが準同型をもつ (k,n) しきい値秘密分散法が医療データを分散バックアップするための条件をすべて満たすため，医療データの分散バックアップに適していると言える．

2.8　まとめ

本章では，初めにバックアップする医療データについて述べた．患者の診察に必要不可欠な電子カルテは災害発生時にすぐ使用することができないため，患者の処方された薬剤情報が乗っているレセプトが災害発生時に使用できるのではないかと注目されている．

次に医療データの分散バックアップに用いる秘密分散法の選定を行った．(k,n) しきい値

表 2.1 (k, n) しきい値秘密分散法と (k, d, n) ランプ型しきい値秘密分散法の比較

分散バックアップ条件	(k, n) しきい値秘密分散法	(k, d, n) ランプ型しきい値秘密分散法
冗長化	○	○
秘匿化	○	△
アクセス制御	○	○
秘匿関数計算	○	△

秘密分散法は，多項式補完を用いてシェアを作成するため，シェアのデータサイズは元データのデータサイズと同じかそれ以上の大きさになる．そこで (k, n) しきい値秘密分散法を用いて作成したシェアの符号効率を改善するために，(k, d, n) ランプ型しきい値秘密分散法が考案された．しかし，(k, d, n) ランプ型しきい値秘密分散法は作成したシェアを $k - t$ 個以上集めることによって，元データの部分的な情報を得ることができるため (k, n) しきい値秘密分散法に比べ秘匿性が低いといえる．(k, d, n) ランプ型しきい値秘密分散法で作成したシェアは秘匿関数計算できない場合があるが，(k, n) しきい値秘密分散法で作成したシェアは秘匿関数計算ができるため，データを復元しなくともデータの演算や解析を行うことができる．よって，医療データの分散バックアップには医療データのの分散バックアップを行うための条件を満たす (k, n) しきい値秘密分散法が適しているといえる．

第 3 章

医療データの部分秘匿化復元

本章では，(k, n) しきい値秘密分散法で分散したデータを部分的に秘匿したまま復元する方法を提案する．

初めに，高知県のバックアップデータの利活用の計画について述べる．次に，災害時に医療データの分散バックアップデータの利活用を述べる．次に，部分秘匿化復元に対する要求を述べ全体構成を述べる．最後に部分秘匿化復元方法，シェアに対するアクセス制限を述べる．

3.1 高知県のバックアップデータの利用計画

本節では，岩手県の医療データが津波によって喪失したことを受けて高知県で進められているバックアップデータの利用計画について述べる．

高知県は医療データの県外バックアップのみならず，バックアップしたデータをリストア以外の用途に利用する計画も立てている．図 3.1 のように平常時に，各医療機関でバックアップした医療データを高知県へき地医療情報ネットワークに参加している医療機関でバックアップした医療データを相互利用する計画や，災害時や救急時に患者の名前や生年月日，保険者番号などを用いてバックアップデータの中から患者の医療データを安全に検索し，仮設診療所やドクターヘリ，ドクターカーなど十分な通信環境が確保できない場所でも利活用できるように，重要度の高いデータから送信かつ短時間で送信できるようにする計画である．

各医療機関での医療データの相互利用が可能になると，患者が他の医療機関に移ることを

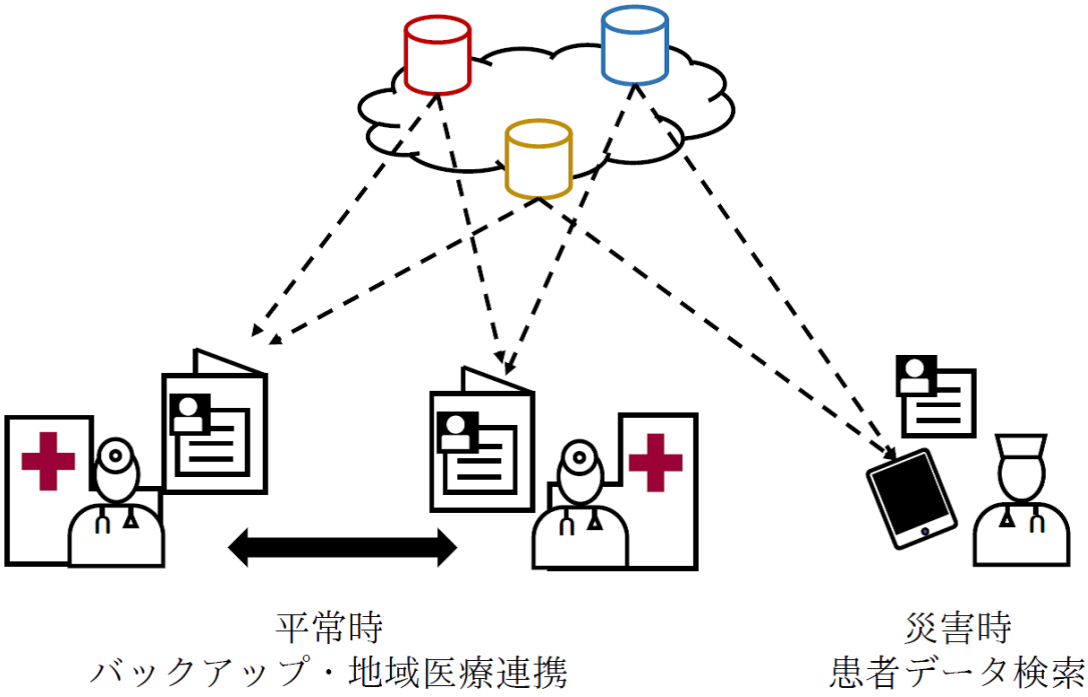

図 3.1　バックアップデータの利活用

スムーズに行うことができる．患者の医療データをバックアップデータの中から検索可能になると，災害時に患者が保険証やお薬手帳などの記録を持っていなかったとしても，患者の名前や生年月日など確認しやすい情報から患者個人の医療データを検索し，患者の病歴や薬剤の服用記録などから医師は患者に対して適切な処置を行うことができる．

このように医療データを利活用することができれば医療分野にとって様々なメリットがある．

しかし，医療データが個人情報を含むデータであるため，図 3.1 のようなシステムを実際に使用するのは，困難である．

3.2　災害時に分散バックアップした医療データの利活用

本節では，(k, n) しきい値秘密分散法を用いて分散バックアップした医療データの災害時における利活用について述べる．

災害時には，負傷者が増え，災害が発生した地域の医療機関も被災している可能性が高

いため，地域の医療機関のみでは負傷者を裁くことは困難である．そこで，災害派遣医療チームである DMAT や，日本医師会に所属する医師や看護師，事務員などで構成された JMAT，日本赤十字社がその地域の医療機関の手助けに入る．

DMAT は，医師，看護師，それ以外の医療従事者で構成され，大規模災害や，負傷者が多い事故などの現場に派遣され，災害や事故発生から 48 時間以内に活動できる機動性を持った専門的な訓練を受けた医療チームのことである．DMAT は，災害発生から 72 時間までの活動を前提としているため，地域の医療機関が完全に復旧していなくとも引き上げることがある．そこで日本医師会に属する医師や看護師の医療従事者で構成された JAMT や日本赤十字社などは，DMAT を引き継いで地域の医療機関の支援を行う．

DMAT や JMAT などの地域外部から派遣された医療従事者は地域の医療機関の復旧や，負傷者のトリアージ，負傷者や被災者のプライマリケアやヘルスケアを行う．その際に，バックアップした医療データが利活用できるとすると，外部から派遣された医療従事者は患者に対して適切な処置を行うため，医療データの閲覧要求を行う．しかし，医療データには個人情報が含まれているため，患者本人の同意がなければ，第三者に医療データを閲覧させることはできない．そこで，図 3.2 のような (k, n) しきい値秘密分散法で分散した医療データを配布できる仕組みがあればよい．データを預けた医療機関の医療従事者からの医療データ要求には，その医療機関の医療データ全てを渡し，外部から派遣された医療従事者からの医療データ要求には，医療機関の医療データから患者の診察に必要なデータのみを渡すような仕組みがあればよい．

しかし，(k, n) しきい値秘密分散法で分散したデータは，復元できるかできないかのどちらかであり，リストアには使用することができるが，一部のデータのみを復元することや，逆に一部のデータを秘匿して復元することはできない．部分的に秘匿したまま復元することができれば，図 3.2 のような仕組みを作成することができる．

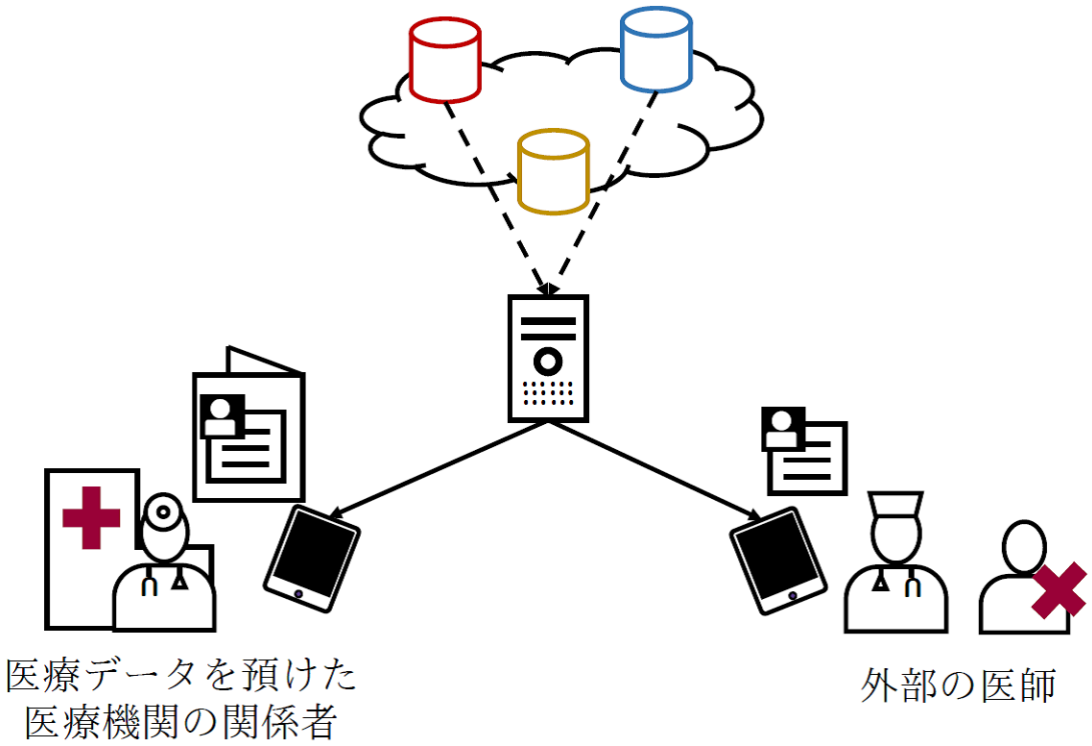

図 3.2　医療データ配布の仕組み

3.3　部分秘匿化復元のための要求

図 3.2 のような仕組みを実現させるには，(k, n) しきい値秘密分散法で分散バックアップしたデータを部分的に秘匿したまま復元する必要がある．部分秘匿化復元を提案するにあたり，部分的に秘匿したまま復元するために必要な要求をまとめる．本節では，部分秘匿化復元に必要な要求を提示する．

図 3.2 を実現するために必要な要求を述べる．まず，分散バックアップデータから，診察に必要なのデータのみを抜き出すことができたとしても，抜き出したデータが正常に復元されていなければ使用することができないため，秘匿部分以外のデータはすべて正常に復元されなければならない．また，診察している患者が変わったときに，医療データも患者個人のデータに変更しなければならないため，秘匿する部分の変更ができるようにしなければならない．

医療データを医療機関ごとに分散バックアップするため，預けた医療機関ごとに医療デー

タのデータの形式が異なる可能性がある．よって，あらゆるデータ形式に対応可能でなければならない．また，各医療機関が自組織の預けた分散バックアップしたデータをリストアして活用できるように，バックアップした医療データを秘匿せずに復元できる必要がある．

部分秘匿化復元を構成するにあたっての要求を以下にまとめる．

要求 1　秘匿部分以外のデータはすべて正常に復元可能であること

要求 2　秘匿部分の変更が可能であること

要求 3　どのようなデータ形式でも対応可能であること

要求 4　部分秘匿化復元だけでなく，データを普通に復元可能であること

以上の前提要求を満たす方法を提案する．

3.4　部分秘匿化の全体構成

図 3.2 のような仕組みを実現させるには (k, n) しきい値秘密分散法で分散バックアップしたデータを部分的に秘匿したまま復元する必要がある．本節では，提案する部分秘匿化復元の全体構成の説明を行う．部分秘匿化復元の構成を図 3.3 に示す．

データを S，S を分割したものを $S_i (i = 1, \cdots, d)$，S_i のシェア集合を $W_i = \{w_{i,1}, \cdots, w_{i,n}\}$，シェア全体を $W = \{W_1, \cdots, W_d\}$，結合ベクトル U_i とする．秘匿部分を $S_s \, (1 \leq s \leq d)$ とし，そのシェア集合を W_s，S_s を秘匿結合する結合ベクトルを U_s とする．

分散時は，まず S を分割する．分割された S_i を (k, n) しきい値秘密分散法で分散し，ネットワーク上にある複数のストレージに保存する．

復元時は，データを復元したいユーザは管理者に U_s を貰い，リファレンスモニタからシェア集合 $W_1, \cdots, W_d$ それぞれのシェアを k 個以上受け取り，U_s を用いてシェア集合を結合する．結合したものを復元すると，S_s が秘匿された S' を得ることができる．

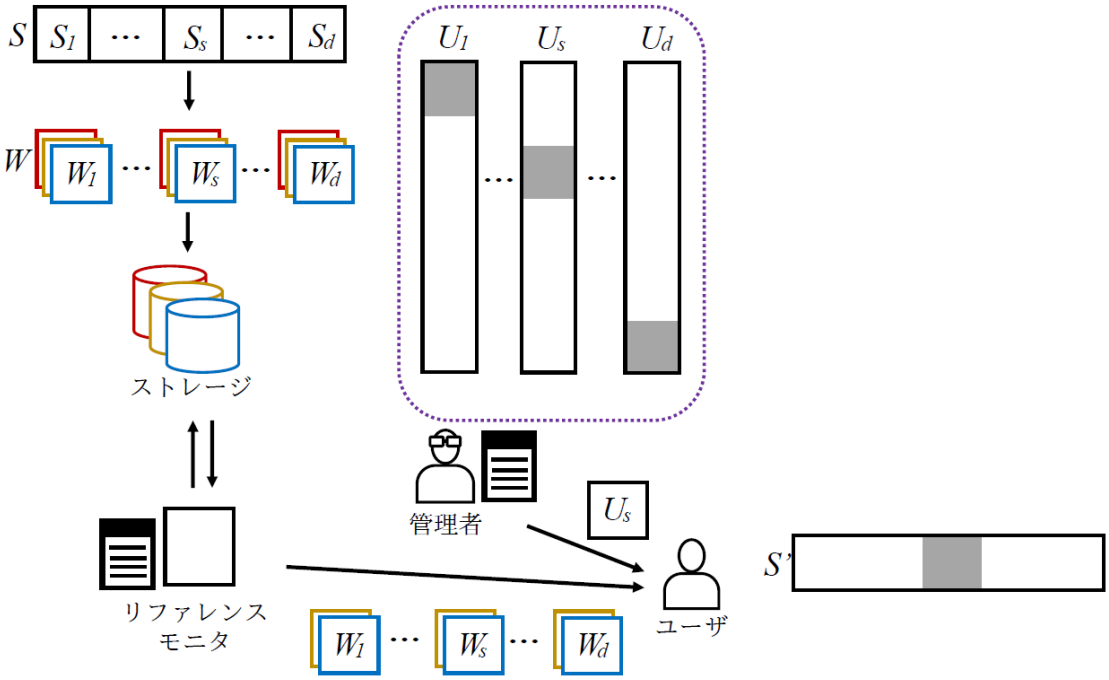

図 3.3　部分秘匿化復元の構成

3.5　部分秘匿化復元

医療データを利活用するには，医療データに個人情報が含まれるため診察に不必要なデータを秘匿して利用しなければならない．しかし，データを (k, n) しきい値秘密分散法で分散すると，データの復元は復元できるかできないかのどちらかである．本節では，前節で提示した要求を満たすような部分秘匿化復元の方法を示す．初めに (k, n) しきい値秘密分散法を用いて部分的に秘匿したまま復元するためのデータの分割方法，結合方法について述べる．次に部分秘匿化復元の詳細な手順を述べる．次に，ダミーデータである α の検証を行う．最後に，提案方法が前節で提示した要求を満たすかどうかを確かめる．

部分的に秘匿したまま復元するために，バックアップするデータを分割する．分割方法は各分割データに分割データサイズ分だけ左シフトして直和をとることで元データになるように分割する．各分割データのデータサイズを l_i とする．式にすると式 3.1 となる．この分割方法ではデータを任意の部分のところで区切ることができる．分割したデータを結合する際に，秘匿するデータ部分にダミーデータを置くことで秘匿データ部分が秘匿されたデータを

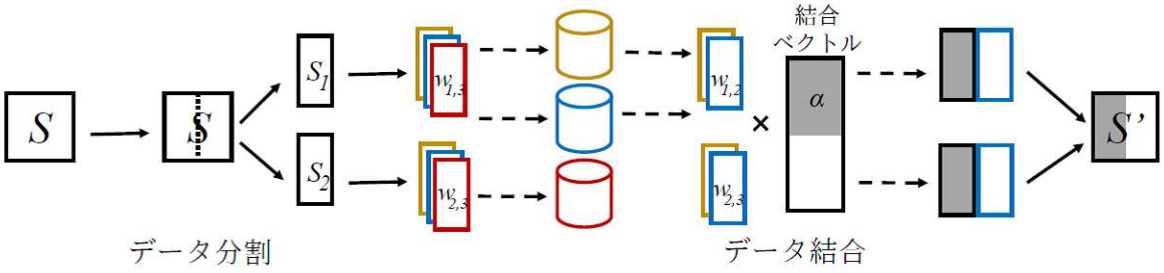

図 3.4　部分秘匿化復元のデータの流れ

得ることができる．データの流れを図 3.4 に示す．

$$S = S_1 \prod_{m=2}^{d} 2^{l_m} + S_2 \prod_{m=3}^{d} 2^{l_m} + \cdots + S_s \prod_{m=s+1}^{d} 2^{l_m} + \cdots + S_{d-1}2^d + S_d \tag{3.1}$$

それぞれの S_i を同じ (k, n) しきい値秘密分散法を用いて分散することで，W_i のシェアを k 個以上集めると S_i を復元することができる．また，(k, n) しきい値秘密分散法は秘匿関数計算が可能なため，W_i のシェアを結合し，結合したシェアを復元することによって S を得ることができる．このため，結合する際に，秘匿するデータ部分である S_s の部分にダミーデータである α を置くことで S_s が秘匿されたデータを得ることができる．

部分秘匿化復元の詳細な手順を以下に示す．秘匿するデータを S_s とする．

1. データ S 式 3.1 を用いて分割し，S_i を作成する．
2. $S < p$ かつ $n < p$ となるような素数 p を選択する．
3. $\mathbb{Z}/p\mathbb{Z} - 0$ の集合 $X_i = \{x_{i,1}, x_{i,2}, \cdots, x_{i,n}\}$ から $n \times k$ の vandermond 行列 $\boldsymbol{X}_i$ を作成する．

$$\boldsymbol{X}_i = \begin{pmatrix} 1 & x_{i,1} & \cdots & x_{i,1}^{k-1} \\ 1 & x_{i,2} & \cdots & x_{i,2}^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{i,n} & \cdots & x_{i,n}^{k-1} \end{pmatrix} \pmod{p} \tag{3.2}$$

4. $\mathbb{Z}/p\mathbb{Z}$ の集合 $R_i = \{r_{i,1} \cdots, r_{i,k-1}\}$(ただし $r_{i,k-1} \in \{\mathbb{Z}/p\mathbb{Z}\} - \{0\}$ とする．) からラ

ンダムに選択し，$(k-1)\times 1$ の乱数行列 $\boldsymbol{R}_i$ を作成する．

$$\boldsymbol{R}_i = \begin{pmatrix} r_{i,1} \\ r_{i,2} \\ \vdots \\ r_{i,k-1} \end{pmatrix} \tag{3.3}$$

5. 式 3.4 を用いて，シェア $w_{i,j} \quad (j = 1, \cdots, n)$ を作成する．

$$\boldsymbol{X}_i \begin{pmatrix} S_1 \\ r_{i,1} \\ \vdots \\ r_{i,k-1} \end{pmatrix} = \begin{pmatrix} w_{i,1} \\ w_{i,2} \\ \vdots \\ w_{i,n} \end{pmatrix} \pmod{p} \tag{3.4}$$

6. $w_{i,j}$ の集合を W_i を

$$W_i = \begin{pmatrix} w_{i,1} \\ w_{i,2} \\ \vdots \\ w_{i,n} \end{pmatrix} \tag{3.5}$$

と定める．

7. W_i をネットワーク上の複数のストレージに分散保管する．

8. 複数のストレージに分散保管されたシェアを集める．取得したシェアからシェア行列 $\boldsymbol{W}_s$ を作成する．

$$\boldsymbol{W}_s = \begin{pmatrix} w_{1,1} & w_{2,1} & \cdots & w_{s,1} & \cdots & w_{d,1} \\ w_{1,2} & w_{2,2} & \cdots & w_{s,2} & \cdots & w_{d,2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{1,k} & w_{2,k} & \cdots & w_{s,k} & \cdots & w_{d,k} \end{pmatrix} \tag{3.6}$$

9. $\boldsymbol{W}_s$ に対して分割データサイズ l_i から作成した結合ベクトル U_s をかける．

$$U_s = \begin{pmatrix} \prod_{m=2}^{d} 2^{l_m} \\ \prod_{m=3}^{d} 2^{l_m} \\ \vdots \\ \alpha \prod_{m=s+1}^{d} 2^{l_m} \\ \vdots \\ 2^d \\ 1 \end{pmatrix} \tag{3.7}$$

$$\begin{pmatrix} w_{1,1} & w_{2,1} & \cdots & w_{s,1} & \cdots & w_{d,1} \\ w_{1,2} & w_{2,2} & \cdots & w_{s,2} & \cdots & w_{d,2} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{1,k} & w_{2,k} & \cdots & w_{s,k} & \cdots & w_{d,k} \end{pmatrix} \begin{pmatrix} \prod_{m=2}^{d} 2^{l_m} \\ \prod_{m=3}^{d} 2^{l_m} \\ \vdots \\ \alpha \prod_{m=s+1}^{d} 2^{l_m} \\ \vdots \\ 2^d \\ 1 \end{pmatrix} \pmod{p} \quad (3.8)$$

これを展開すると

$$\boldsymbol{X}_i \begin{pmatrix} S_1 \prod_{m=2}^{d} 2^{l_m} + S_2 \prod_{m=3}^{d} 2^{l_m} + & \cdots & +\beta \prod_{m=s+1}^{d} 2^{l_m} + & \cdots & +S_{d-1}2^d + S_d \\ r_{1,1} \prod_{m=2}^{d} 2^{l_m} + r_{2,1} \prod_{m=3}^{d} 2^{l_m} + & \cdots & +\beta \prod_{m=s+1}^{d} 2^{l_m} + & \cdots & +r_{d-1,1}2^d + r_{d,1} \\ & \vdots & & \vdots & \\ r_{1,k-1} \prod_{m=2}^{d} 2^{l_m} + r_{2,k-1} \prod_{m=3}^{d} 2^{l_m} + & \cdots & +\beta \prod_{m=s+1}^{d} 2^{l_m} + & \cdots & +r_{d-1,k-1}2^d + r_{d,k-1} \end{pmatrix}$$

となり，復元すると

$$S_1 \prod_{m=2}^{d} 2^{l_m} + S_2 \prod_{m=3}^{d} 2^{l_m} + \cdots + \beta \prod_{m=s+1}^{d} 2^{l_m} + \cdots + S_{d-1}2^d + S_d = S' \quad (3.9)$$

となる．β は α を復元した結果である．S_s の部分が β に置き換わるため，S_s のデータを復元することができない．よって，部分的にデータを秘匿して復元することができる．

ダミーデータである α が 0 の場合と α が 0 以外の場合の検証を行う．部分秘匿化復元が可能か，結合したシェアから秘匿部分が分かるかを α を変えて検証した．検証結果を表 3.1 に示す．

α が 0 の場合，復元データは秘匿部分以外正常に復元されていることが確認できた．しかし，結合ベクトルを用いてシェアを結合したものを見ると，秘匿部分が 0 になり，どこの部分が秘匿されているかすぐに分かる．よって 0 の部分のシェアが，秘匿部分のシェアであると分かり，その部分のシェアを何らかの方法で入手することによって，データを不正に復元される可能性がある．

α が 0 以外の場合について，$\mathbb{Z}/p\mathbb{Z}$ 上で適当に選択した a，$\mathbb{Z}/p\mathbb{Z}$ 上で適当に発生させた乱数列 $R = \{r_1, \cdots, r_n\}$，R を分散データと同じ (k, n) しきい値秘密分散法で作成したシェ

表 3.1　α の検証結果

α	部分秘匿化復元可能か	秘匿部分判定不可能か
0	○	×
a	○	×
R	×	○
W_R	○	○

ア $W_R = \{w_{R,1}, \cdots, w_{R,n}\}$ の 3 パターンの検証を行う．

はじめに，$\mathbb{Z}/p\mathbb{Z}$ 上で適当に選択した a を α として試した．結果，α が 0 の場合と同じ結果になった．次に，$\mathbb{Z}/p\mathbb{Z}$ 上で適当に発生させた乱数列 R を α として試した．結果，復元データは元データとは完全に異なるデータになることが確認された．最後に適当に発生させた乱数列 R のシェア W_R を α として試した．結果，部分的に秘匿されたデータを得ることができ，シェアを結合したものを見ても秘匿部分が分からないことが確認できた．しかし，データごとに乱数列と乱数列のシェアを作成保管しなければならないため，ストレージに保持するデータ数が増える．

よって，W_R が，結合したシェアから秘匿部分を割り出されないため，α に適していると言えるが，保管するシェアの個数が増えるため，本提案では，α に 0 を用いる．

提案方法が前節で示した要求を満たすかどうかを確認する．要求 1 の秘匿部分以外のデータはすべて復元できることは式 3.9 から分かる．要求 2 の秘匿部分の変更が可能であることは，U_s の α を置く場所を変更することで秘匿部分の変更が可能である．また α を置く場所を増やすことで秘匿部分を広げることも可能である．要求 3 のどのようなデータ形式でも対応可能であることは，データを分割する場所を変更することで可能である．要求 4 の部分秘匿化復元だけでなく，データを普通に復元可能であることは，結合ベクトルに α を置かけなければ，式 3.1 となり，データを普通に復元することができる．

よって，提案方法は前節で提示した要求 4 つをすべて満たす．

3.6　シェアへのアクセス制限方法

前節で示した部分秘匿化復元では，何らかの方法で W にアクセスすることができれば，l_i から作成できる結合ベクトルと組み合わせて S を復元される可能性がある．よってシェアに対してアクセス制限をかける必要がある．本節では，シェアに対してのアクセス制限方法を述べる．

W へのアクセス制限をかけるために，W に唯一アクセスできるリファレンスモニタと U_i を作成する管理者を設置する．W を保持している複数のストレージは，リファレンスモニタ以外のユーザがアクセスしてきた場合，要求を破棄するようにする．そのため，ユーザがシェアを取得する際にはリファレンスモニタに問い合わせるようにする．

U_i は不正に作成される恐れがあるため，U_i の作成を管理者に限定する．管理者が作成した U_i にはラベルをつけ U_i' とする．このラベルの情報をリファレンスモニタに送る．

リファレンスモニタは，W_i を含まないシェア群 $\boldsymbol{G_i'} = \{W_1, \cdots, W_{i-1}, W_{i+1}, \cdots, W_d\}$ を作成し，U_i' と対応するように管理者から受け取ったラベル情報を用いて対応付ける．対応付けた情報をラベル対応表とし，保管する．

要求を受けたリファレンスモニタは，まずユーザが U_i' を持っているか確認し、持っていなければ要求を破棄する．次に，U_i' につけられたラベルと同様のラベルがついた $\boldsymbol{G_i'}$ をラベル対応表から探し，$\boldsymbol{G_i'}$ に属する $W_1, \cdots, W_{i-1}, W_{i+1}, \cdots, W_d$ からデータが復元できるようにシェアを k 個以上集めユーザにわたす．

このように管理者とリファレンスモニタをおくことで，データを復元したいユーザはまず管理者に U_i の取得要求を送り，管理者から受け取った U_i' を用いてリファレンスモニタにシェアの要求を送ることになる．よって，全てのユーザはデータを復元する際に管理者から U_i' をもらう必要があるため，管理者の許可のない第三者が S を復元することを防ぐ．

しかし，作成した $\boldsymbol{G_i'}$ は多元的な構造を持つ．例えば，ユーザが，$\boldsymbol{G_i'}$ と $\boldsymbol{G_i'}$ 以外の $\boldsymbol{G_b'}$ を保持していた場合，$\boldsymbol{G_i'}$ と $\boldsymbol{G_b'}$ を組み合わせて W_i を入手することができる．よって $\boldsymbol{G_i'}$ に対して多元的なアクセス制限をかける必要がある．

そこで，ユーザが入手できる U_i' を一つに限定する．管理者はユーザに対して U_i' を配布するとき，一つだけ配布するようにし，管理者はユーザがどの U_i' を保持しているかの記録をつける．記録をつけるものを配布記録とする．配布記録にはユーザと U_i' が 1 対 1 になるように記録し，ユーザがどの U_i' を持っているかどうかを把握する．一つ以上の U_i' を所持しているユーザが U_i' 以外の U_b' 取得要求を送ってきた場合，管理者はその要求を破棄する．このように管理者は配布記録を用いてユーザが複数の U_i' を持つことを防ぐ．

シェアへのアクセス手順を示す．秘匿するデータを S_s として説明を行う．アクセス手順の流れを図 3.5 に示す．

1. ユーザが管理者に対して，U_s の取得要求を送る．
2. 要求を受けた管理者は要求を送ってきたユーザが U_s' 以外の U_i' を持っていないか配布記録でチェックし，持っていなければユーザに U_s' を配布し，配布記録に記録する．持っている場合は要求を破棄する．
3. ユーザは取得した U_s' を用いてリファレンスモニタにシェアの要求を送る．
4. リファレンスモニタはユーザが U_s' を持っているかを確認し，持っていなければ要求を破棄する．
5. リファレンスモニタはユーザが送ってきた U_s' にラベルがついているかを確認する．ついていなければ要求を破棄する．
6. リファレンスモニタは，ユーザの U_s' につけられているラベルと同様のラベルがついた $\boldsymbol{G}_s'$ に属しているシェアの集合 $W_1, \cdots, W_{s-1}, W_{s+1}, \cdots, W_d$ からデータを復元できるようにシェアを k 個以上集めユーザにわたす．

以上のアクセス制限により，ユーザが不正な結合ベクトルでシェアの要求を送ってきた場合，リファレンスモニタが要求を破棄し，ユーザが U_b' を保持したまま U_i' を要求してきた場合，管理者が要求を破棄するため，ユーザが管理者の意図しないシェアに対してのアクセスを防ぐことができる．

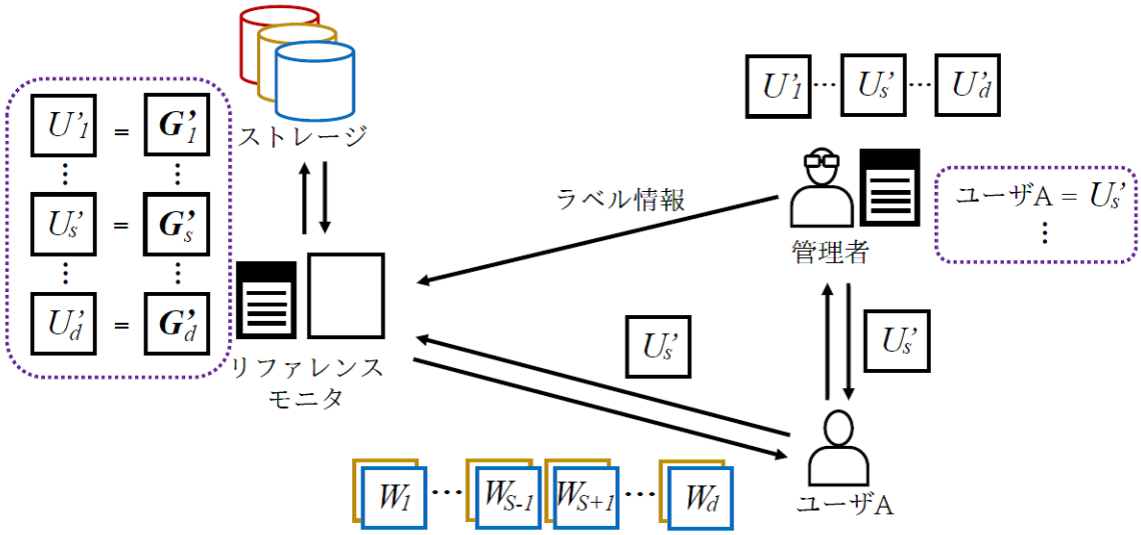

図 3.5　アクセス手順の流れ

3.7　まとめ

本章では，まず高知県のバックアップデータを用いた利活用の計画について述べた．高知県へき地医療情報ネットワークに参加している病院間で医療データのやりとりを行うことや，災害時や救急時に確認しやすい患者の名前や生年月日から患者個人の医療データを探し出して活用できるようになれば，地域医療連携など医療分野にとって様々なメリットがある．しかし，医療データには個人情報が含まれているため，安易に利活用することはできない．次に，(k, n) しきい値秘密分散法で分散バックアップした医療データを災害時に利活用することについて述べた．災害時に，分散バックアップデータから DMAT など外部から派遣された医師が患者の診察に必要なのデータのみを閲覧することができれば，患者に対して的確な処置を行うことができるが，(k, n) しきい値秘密分散法を用いて分散したデータを部分的に復元することや部分的に秘匿したまま復元することはできない．部分的に秘匿したまま復元することが可能であれば，災害時に利活用することができる．

そこで，データを部分的に秘匿したまま復元する方法を示した．データをそのまま (k, n) しきい値秘密分散法を用いてシェアを作成すると，データを部分的に秘匿したまま復元することができないため，データを分割した．分割したデータごとに同じ (k, n) しきい値秘密分

散法を用いてシェアを作成し，作成したシェアを秘匿する部分にダミーデータ α が置かれた結合ベクトルを用いて結合し，復元すると部分的に秘匿されたデータを得ることができる．しかし，提案方法では，何らかの方法でシェアにアクセスすることができれば，結合ベクトルを用いてシェアを結合し，データを不正に復元される恐れがある．そのため，シェアに対してアクセス制限方法を示した．シェアにアクセス制限をかけるために，シェアに唯一アクセスできるリファレンスモニタと，結合ベクトルを作成する管理者を置いた．結合ベクトルは不正に作成される恐れがあるため，管理者が作成した結合ベクトルにはラベルをつけた．リファレンスモニタは，データを復元したいユーザが管理者が作成したラベル付きの結合ベクトルを持っているか確認し，持っていなければシェアを配布しない．そのため，管理者の意図しないユーザにシェアを取得されることを防ぐことができる．

部分秘匿化復元とシェアへのアクセス制限を用いることによって，(k, n) しきい値秘密分散方で分散したバックアップデータから診察に必要なの医療データを抜き出すことができるといえる．

第 4 章

部分秘匿化復元とシェアに対するアクセス制限の評価

本章では，前章で示した部分秘匿化復元とシェアに対するアクセス制限の評価を行う．初めに (k, n) しきい値秘密分散法を用いた部分秘匿化復元の評価を行い，次にシェアへのアクセス制限についての評価を行う．最後に全体の考察を行う．

4.1 (k, n) しきい値秘密分散法を用いた部分秘匿化復元の評価

本節では，(k, n) しきい値秘密分散法を用いた部分秘匿化復元の評価を行う．データの分割についての評価を行い，その後にシェアの結合についての評価を行う．

4.1.1 データ分割の評価

本提案では，部分的にデータを秘匿して復元するために，データを分割し，分割データごとに同じ (k, n) しきい値秘密分散法を用いて分割データのシェアを作成する．

データの分割には式 4.1 を用いる．S はデータ，$S_1, \cdots, S_d$ は各分割データ，$l_i (i = 1, \cdots, d)$ は分割データサイズである．

$$S = S_1 \prod_{m=2}^{d} 2^{l_m} + S_2 \prod_{m=3}^{d} 2^{l_m} + \cdots + S_s \prod_{m=s+1}^{d} 2^{l_m} + \cdots + S_{d-1} 2^d + S_d \qquad (4.1)$$

この分割方法では l_i の値を変えることによってデータを任意の部分で区切ることができる．データを任意のところで区切ることができるため，どのようなデータ形式でも対応すること

表 4.1 (k, n) しきい値秘密分散法を用いて作成したシェアの個数比較

(k, n) しきい値秘密分散法を適応するデータに対する処理	何もしない	分割
作成するシェアの個数	n	dn
復元可能なシェアの個数	k	dk

ができると言える．よって，標準規格が定まっていない電子カルテや発行する医療機関によって異なるレセプトに対して有効な手段であるといえる．また，データを分割することにより分割データに重要度をつけることができる．分割データにつけられた重要度をシェアにも持たせ，重要度順にストレージから取り出すことで，重要度ごとにデータを送信することができる．

表 4.1 は，データをそのまま (k, n) しきい値秘密分散法を用いてシェアを作成したときと，データを分割し分割データそれぞれを (k, n) しきい値秘密分散法を用いてシェアを作成したときのシェア個数とデータを復元するのに必要なシェアの個数を示している．データを分割したことによって，管理するシェアの個数が n 個から dn 個に増加した．また，分割データごとにシェアが存在するため，どれがどの分割データのシェアなのかを識別子をつけるなどの方法で区別しなければならない．(k, n) しきい値秘密分散法は任意のシェアを k 個以上集めることによって，元データを復元することができるが，データを分割することで，分割データごとにシェアを k 個以上集めなければならず，元データを復元する場合は dk 個以上集めなければならない．よって容易にシェアを集めることが困難になった．

データを分割することによって，あらゆるデータ形式に対応でき，分割データごとに重要度をつけることができるが，シェアの個数が増え管理を容易に行うことが困難になったと言える．

4.1.2 シェアの結合についての評価

秘匿したい分割データのシェア部分にダミーデータ α を置き，結合時にマスキングを行い結合することでデータを部分的に秘匿したまま復元する．結合には，分割データのデータサ

イズ l_i から作成した式 4.2 の結合ベクトルを使用する．

$$U_i = \begin{pmatrix} \prod_{m=2}^{d} 2^{l_m} \\ \prod_{m=3}^{d} 2^{l_m} \\ \vdots \\ \alpha \prod_{m=i+1}^{d} 2^{l_m} \\ \vdots \\ 2^d \\ 1 \end{pmatrix} \tag{4.2}$$

l_i が分かればあらゆるデータ形式に対応した結合ベクトルを作成することができる．また，式 4.2 の α を置く場所は任意に変更することができる．よって秘匿部分を変更することができる．

データ形式や秘匿部分の変更に対応するために，それぞれに対応した結合ベクトルを作成しなければならないため，結合ベクトルの数が増える．データ形式のパターン数が n 個，秘匿部分のパターン数が m 個とすると結合ベクトルの個数は $n \times m$ となる．

ダミーデータである α は本提案では 0 を用いている．0 は結合したシェアを見ると秘匿部分が 0 になるため，どこを秘匿しているかが容易に分かる．よって，その部分のシェアを何らかの方法で入手することができればデータを不正に復元される可能性がある．シェアを結合したものから秘匿部分を分からないようにするためには，適当に発生させた乱数列を分割データと同じ (k, n) しきい値秘密分散法で作成したシェアを α にすればよい．しかし，分割データごとに乱数列と乱数列のシェアを保持しなければならない．そのため乱数列一つのデータサイズが R とすると，乱数列のシェア全体のデータサイズは nR となり，保持するデータ数が増える．α を 0 にするとシェアを結合したものから秘匿部分が分かる．シェアを結合したものから秘匿部分が分からないようにするためには乱数列のシェアが適しているが，全体のデータ数が増える．よってダミーデータである α の選定は困難である．

提案方法のシェアの結合は，結合ベクトルのダミーデータ α を置く場所を変更することで秘匿部分の変更することができる．また，各分割データのデータサイズ l_i が分かれば色々なデータ形式にも対応することができる．α や l_i を変更することによって秘匿部分の変更

やデータ形式の対応ができるが，それぞれのデータ形式に対応する結合ベクトルやそれぞれの秘匿部分に対応する結合ベクトルを作成しなければならないため，結合ベクトルの数が増える．

4.2　シェアへのアクセス制限についての評価

提案した部分秘匿化復元は，何らかの方法でシェアを入手することができれば，シェアと結合ベクトルを用いてデータを不正に復元される恐れがある．そのため，シェアにアクセス制限をかけなければならない．本節では，シェアへのアクセス制限についての評価を行う．

シェアにアクセス制限をかけるために，シェアへのアクセスをリファレンスモニタに限定し，さらに U_i が不正に作成される恐れがあるため，U_i を唯一作成できる管理者をおいた．管理者が作成した U_i にはラベルをつけて U_i' とした．リファレンスモニタは管理者が配布する U_i' を持っていないユーザに対してシェアの配布を行わないため，管理者の意図しないユーザがシェアにアクセスすることを防ぐことができる．

ユーザにデータを不正に復元されないよう各分割データのシェア集合 $W_1, \cdots, W_d$ から秘匿する部分のシェア集合 W_s を抜いた $\boldsymbol{G}_s' = \{W_1, \cdots, W_{s-1}, W_{s+1}, \cdots, W_d\}$ を作成した．作成したことにより，シェアの構造が複雑になった．よって，データを復元したいユーザがシェアを入手し部分的に秘匿されたデータを手に入れるまでの段階が増えた．

シェアへのアクセス権や U_i' の作成権を限定することにより，災害時に権限を持ったリファレンスモニタや管理者が被災した場合，シェアにアクセスすることができなくなり，データを復元することができなくなる可能性がある．また，データを復元したいユーザ全員が，同じリファレンスモニタや管理者に要求を送るため，ボトルネックになることが想定される．

本提案のアクセス制限は，単一ユーザに対してのアクセス制限のため，複数のユーザが結託した場合，図 4.1 のようにそれぞれのユーザが保持している U_i' で取得できるシェアを組み合わせて，データを不正に復元される可能性がある．また，U_i' と $\boldsymbol{G}_i'$ を対応させるために

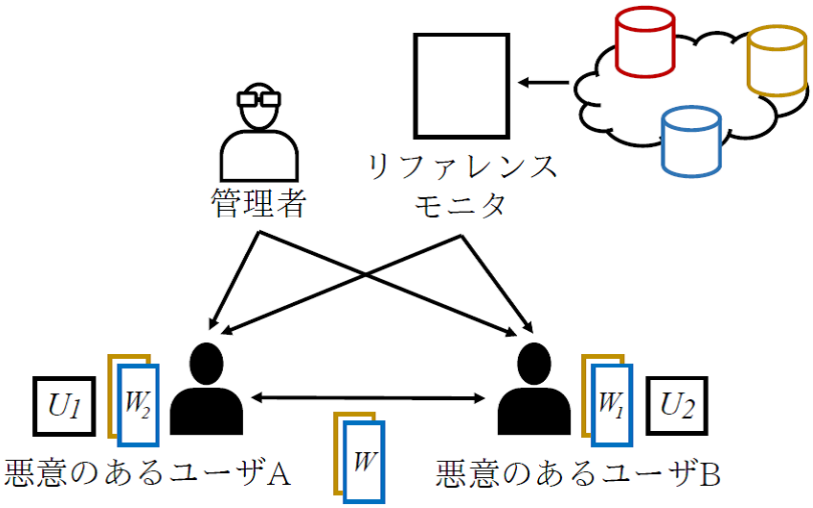

図 4.1　ユーザ同士の結託

使用するラベルの情報が流出すると第三者が不正に U_i' を作成することができる．よって，リファレンスモニタから任意のシェアを取得することができ，データを不正に復元される可能性がある．ラベルの情報が流出しなかったとしても何らかの方法でラベル情報を偽装されれば U_i' に対応する正しい $\boldsymbol{G_i'}$ を入手することができなくなる．

シェアに対するアクセス制限は，管理者の意図しないユーザにデータを不正に復元されることを防ぐことはできるが，ラベル情報の漏えいや，複数人の結託に弱く，不十分なものであるといえる．

4.3　部分秘匿化復元とシェアに対するアクセス制限の考察

(k, n) しきい値秘密分散法のデータの復元は復元できるかできないかのどちらかであり，部分的に秘匿したまま復元することは不可能である．そこで本提案では，データを分割し，分割データそれぞれに同じ (k, n) しきい値秘密分散法を用いてシェアを作る．秘匿するデータ部分にダミーデータを置いた結合ベクトルを用いてシェアを結合し，復元することにより，データを部分的に秘匿して復元している．また，この部分秘匿化復元では，何らかの方法でシェアにアクセスすることができれば，結合ベクトルを用いてデータを不正に復元される恐れがあるため，作成したシェアに対してアクセス制限をかける必要がある．本節では，

4.3 部分秘匿化復元とシェアに対するアクセス制限の考察

部分秘匿化復元とシェアに対しての考察を行う．

提案方法のメリットは，データを任意の部分で分割し，あらゆるデータ形式でも対応することができること，結合ベクトルのダミーデータを置く場所を変更することによって秘匿部分を任意に変更することができること，アクセス制限により管理者の意図しないユーザがデータを不正に復元することを防ぐことができることである．

一方，デメリットは，データを分割したことによりシェアの管理が複雑になったこと，アクセス制限によって，シェアを入手するまでの段階が増えたこと，全てのユーザが管理者とリファレンスモニタに要求を送るため，管理者とリファレンスモニタがボトルネックになること，複数のユーザの結託に弱いことである．

シェア入手までの複雑な段階をソフトウェア化することによってユーザがデータを取得するまでの負担を軽減できると考える．まず，図 4.2 のように前節でユーザとしていた役割をタブレットなどの電子端末に持たせる．電子端末には，シェアと U_i' を入手し U_i' を用いてシェアを結合して復元するまでを行うソフトウェアが入っている．ユーザが部分的に秘匿されたデータ S' を入手するまでの流れを以下に示す．

1. ユーザは電子端末に対して S' の要求を行う．
2. 要求を受けた電子端末は管理者に対して U_i' の要求を送る．
3. 管理者は要求を送ってきたタブレットの識別番号を確認し，配布記録を用いて他の U_i' を持っていないかを確認する．持っていなければ U_i' を配布する．
4. 管理者から受け取った U_i' を用いてリファレンスモニタにシェアの要求を行う．
5. リファレンスモニタは U_i' のラベルに対応した $\boldsymbol{G_i'}$ に属するシェアの集合それぞれからシェアを k 以上集め，タブレットに渡す．
6. 電子端末は U_i' を用いてシェアを結合し S' を復元する．
7. 電子端末は復元した S' をユーザに渡す．

このように前章でユーザとしていた役割をタブレットなどの電子端末に持たせることで，ユーザ自身が複雑な手順を踏まなくとも部分的に秘匿されたデータを得ることができる．ま

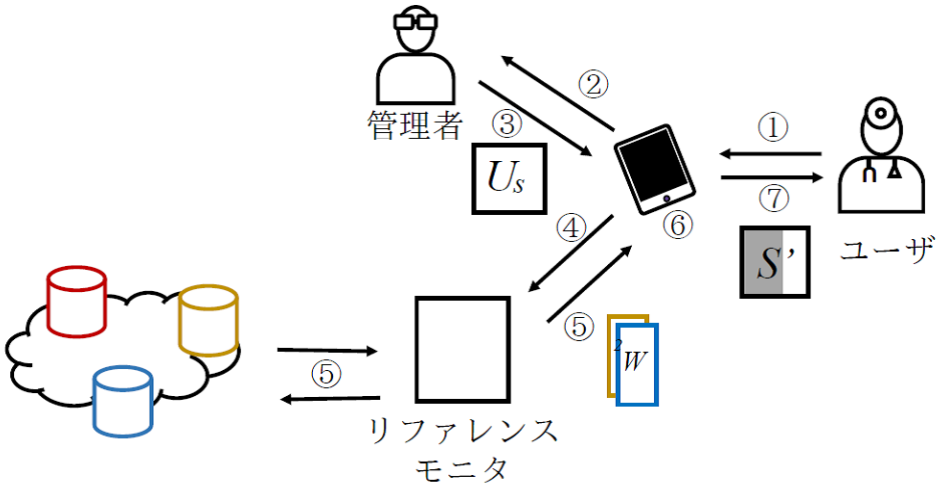

図 4.2　ソフトウェアを用いたシェアと結合ベクトルの取得

た，ソフトウェアをインストールできる電子端末を限定し，特定の電子端末の識別番号を登録し識別番号以外のアクセスを制限すると，特定の電子端末以外の通信を遮断することができるため，情報漏えいのリスクが下がる．災害時に医療データを部分的に秘匿したまま復元することができる電子端末を渡されても使用することが難しいため，地域医療連携など平常時から使用してもらい，ユーザに電子端末の操作に慣れてもらうことで災害時スムーズに活用することができる．

リファレンスモニタと管理者がボトルネックになることは，リファレンスモニタと管理者を冗長化することによって解決できると考える．しかし，リファレンスモニタ，管理者を一つ以上置くことを許可することで，不正なリファレンスモニタ，管理者が現れる可能性がある．そこで，全体の管理者を新たに用意する．全体の管理者は，シェアにアクセスする権限，結合ベクトルを作成する権限を発行管理し，リファレンスモニタと管理者を監視する．全体の管理者が発行する権限が正当なものだと示すために，全体の管理者が発行する権限に電子署名をつける．図 4.3 のように全体の管理者がそれぞれの権限を渡すことによって，リファレンスモニタや管理者になることができる．逆に，全体の管理者の署名がついた権限を持っていなければリファレンスモニタや管理者にはなれない．よって不正なリファレンスモニタや不正な管理者が現れることを防ぐ．また，リファレンスモニタや管理者は破損や故障

4.3　部分秘匿化復元とシェアに対するアクセス制限の考察

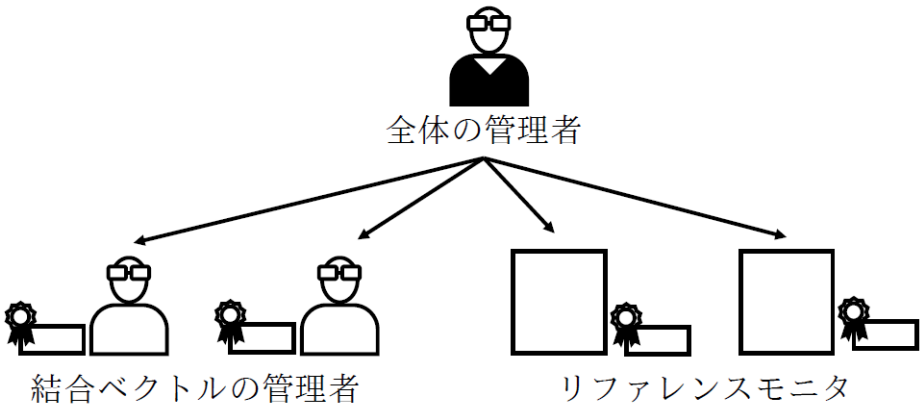

図 4.3　全体の管理者による権限譲渡

によって機能停止する可能性がある．全体の管理者は，リファレンスモニタや管理者が正常に稼働しているかどうかを監視し，機能停止していた場合図 4.4 のように機能停止しているリファレンスモニタや管理者から権限を剥奪し，他のマシンに権限を渡す．権限を剥奪し他のマシンに権限を渡すまでの流れを図 4.5 に示す．また，流れを以下に示す．

1. 全体の管理者は権限を与えたリファレンスモニタや管理者が正常に活動しているか確認するため，生存確認を送る．
2. 生存確認を受けたリファレンスモニタや管理者は生存確認に応答する．
3. 応答しなかったリファレンスモニタや管理者がいる場合全体の管理者は応答しなかったリファレンスモニタや管理者に再度生存確認を送る．
4. 3 回生存確認を送り応答がなければ応答がなかったリファレンスモニタや管理者の権限を剥奪する．
5. 全体の管理者は他のマシンに剥奪した権限を与える．

このようにすることでリファレンスモニタや管理者が故障した場合にも対応することができる．

複数のユーザの結託に弱いため，結託に対して何らかの対策をして使用しなければならない．例えば，図 4.6 のようにデータを復元できる特定のタブレットなどの電子端末を用意

4.3　部分秘匿化復元とシェアに対するアクセス制限の考察

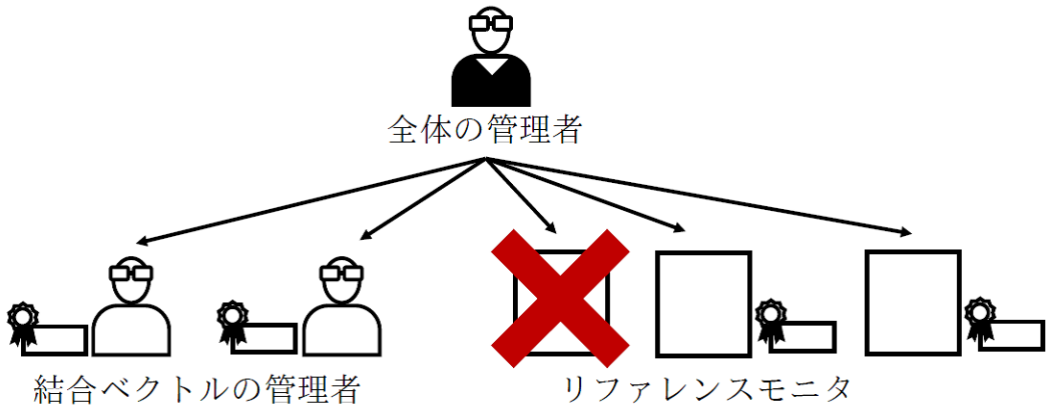

図 4.4　機能停止時の対応

し，ユーザに配布する．ユーザに配布した電子端末間の通信を制限することによって，それぞれの電子端末で入手可能なシェアを交換し，お互いに不足しているシェアを組み合わせて，不正にデータを復元されることを防ぐことができる．

提案方法で部分秘匿化復元を行うことが可能になったが，シェアの管理が複雑になったこ

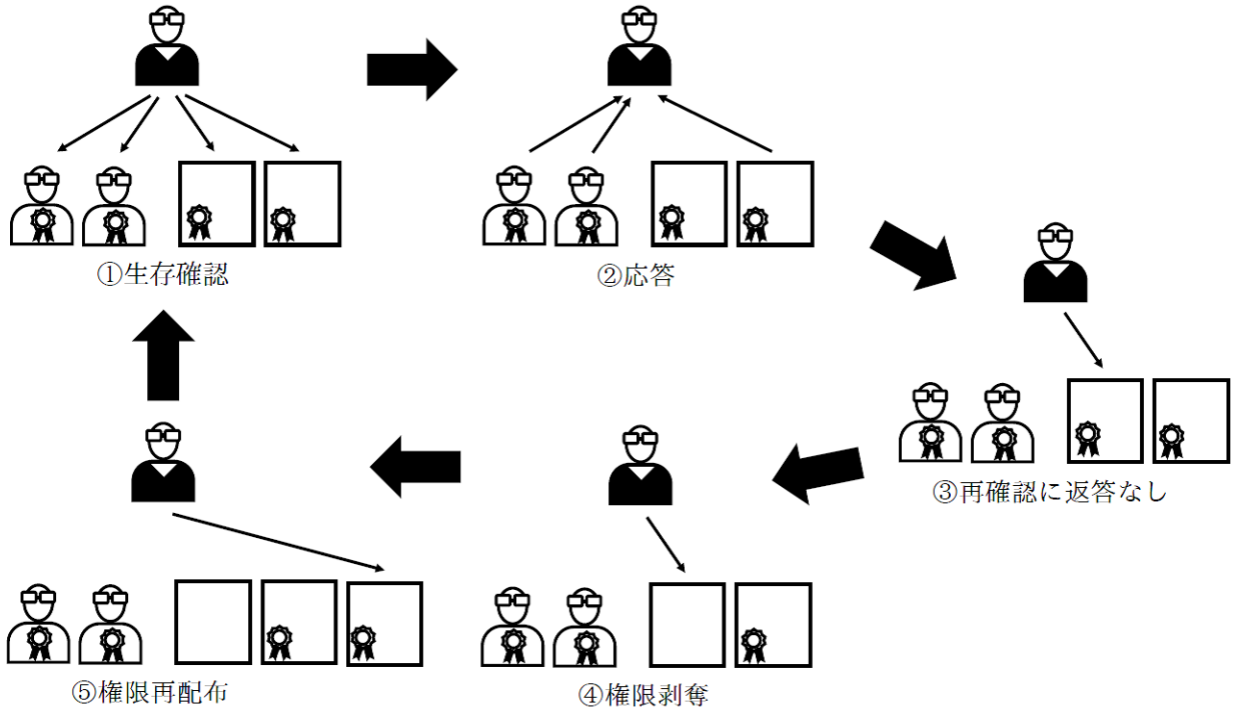

図 4.5　全体の権限剥奪の流れ

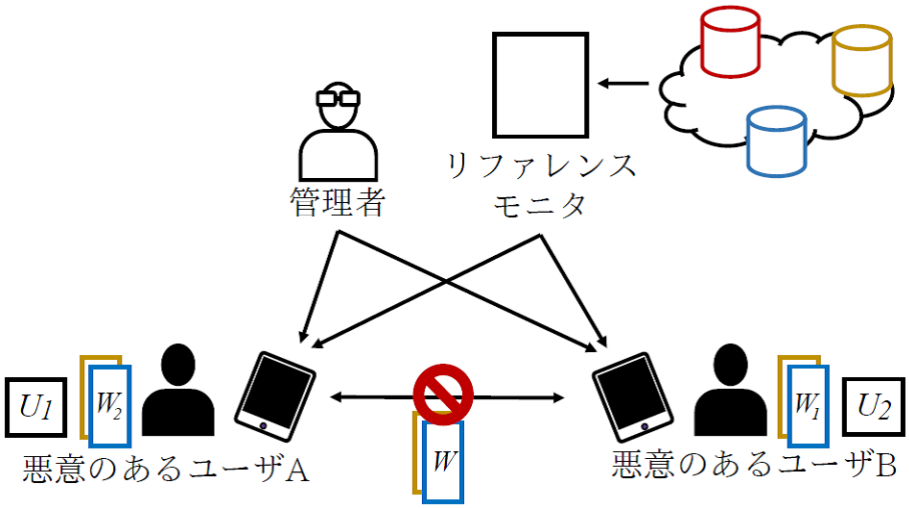

図 4.6　端末間通信の制限

とや，シェアにかけたアクセス制限は複数人の結託に弱いと言うデメリットもある．しかし，デメリットを把握した上で本節で示したような対策を施すことで十分活用可能だといえる．

第 5 章

結論

5.1 本研究のまとめ

本論文では (k, n) しきい値秘密分散法で分散したデータを利活用するために，(k, n) しきい値秘密分散法で分散されたデータを部分的に秘匿したまま復元する方法を示し，提案方法の評価を行った．

提案方法では，データを分割し，それぞれの分割データを同じ (k, n) しきい値秘密分散法で分散した．復元時にそれぞれの分割データのシェアを復元できるように k 個以上集め，秘匿部分にダミーデータを置いた結合ベクトルを用いてシェアを結合し，復元することによって，部分的に秘匿されたデータを得ることができた．

しかし，何らかの方法でシェアを入手することができれば，結合ベクトルは分割データのデータサイズが判明すると容易に作成することができるため，データを不正に復元される可能性があった．そこでシェアに対してアクセス制限方法を示した．アクセス制限によって，管理者の意図しないユーザがシェアにアクセスできないようにした．

提案方法を評価した結果，部分的に秘匿したまま復元するために行ったデータの分割は，分割する部分を任意で決めることができるため，あらゆるデータ形式に対応することができ，分割データに重要度をつけることができる．しかし，分割データごとにシェアを作成するため，シェアの管理や構造が複雑になり，ユーザが容易にシェアを取得することが困難になった．また，結合ベクトルを用いてシェアを結合し，ダミーデータを置く部分を変更することで秘匿する部分を変更することができる．しかし，データ形式や秘匿部分に対応した結合ベクトルを作成する必要があり，結合ベクトルの数が増え，管理が複雑になることが分

かった．

シェアに対するアクセス制限を評価した結果，管理者の意図しないユーザに対してシェアを取得することを防ぐことが可能であると分かった．しかし，シェアに対してのアクセス制限は，単一ユーザに対してのアクセス制限であるため，複数のユーザが結託することによって，データを不正に復元される可能性があることが分かった．

5.2　今後の課題

今後の課題として，提案方法ではシェアの数が増えシェアの構造や管理が複雑になったため，効率的なシェアの管理方法を検討する必要がある．また，結合したシェアから秘匿部分がわからないようなダミーデータ α を検討しなければならない．

シェアに対してのアクセス制限は，複数のユーザが結託することに対して弱いため，複数のユーザが結託できないようにする仕組みや，結託してもシェアを組み合わせることができない仕組みを考える必要がある．

謝辞

本研究を遂行するにあたり，終始ご指導並びにご鞭撻を賜わり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．本卒業研究の副査をしていただいた情報学群植田和憲講師，横山和俊教授のお二人に謹んで感謝致します．また，NOC の職員であり研究室の OB でもある福冨英次氏，修士課程の横田優佳氏にも謹んで感謝致します．

福本教授には，呆れながらも理解力の乏しい私に対して何回もご指導してくださったこと感謝しています．また，いまいち理解せずによく分からない方向に突っ走っていく私の舵をとりながらここまで導いていただいたこと大変感謝しています．

仕事で忙しい中，何かと気にかけてくださった福冨氏には大変感謝しています．福冨氏のアドバイスがなければ本研究がまったく進まなかったと言っても過言ではありません．色々なことをアドバイスしていただいているにも関わらず，思考停止してしまい大変申し訳ありません．

横田氏の専攻とはかけ離れていたのにもかかわらず指導をしてくださった横田氏には大変感謝しています．何回も私の読みにくい文章を読み日本語の指導をしていただいたこと大変感謝しています．横田氏のアドバイスは的をえているにも関わらずいつも聞かなくて大変申し訳ありません．

先生や先輩から投げられた仕事をそのまま投げても，文句一つ言わずこなしてくれた宮西寛奈氏，佐藤諒氏，星野浩希氏には大変感謝しています．特に宮西氏には研究室イベントの仕事をよく投げていました．宮西氏がいなければ成功していなかったイベントはたくさんあったと思います．私は思いつきで行動する上に気分屋なので無茶ぶりをすることが多かったと思いますが，うまいこと受け流してくださってありがとうございます．おかげで楽しい研究室生活を送ることができました．論文を書いている際に飽きたなど皆様の士気を下げるようなことを言ってしまって申し訳ありません．

最後になりましたが，この 4 年間私を支えてくださった皆様に感謝致します．

参考文献

[1] 澤田努, “南海トラフ地震に向けて医療情報を県外保全する取り組みについて 高知県医療通信技術 (ICT) 連絡協議会の立ち上げについて,” 医療情報学, Vol.33, No.4, pp.255–233, 2013.

[2] 福本昌弘，“南海トラフ巨大地震に対する医療情報の保全のための高知県での取り組み,”Mercato, 東北情報通信懇談会, Vol.89, pp.18–20, 2014.

[3] A. Shamir，“How to Share a Secret,” Communication of the ACM, Vol.22, No.11, pp.612–613, Nov. 1979.

[4] 舟橋釈仁，“秘密分散法を用いた広域ファイルシステム,” 平成 14 年度高知工科大学修士学位論文， Feb．2003.

[5] 黒田知宏，木村映善，松村奏志，山下芳範，平松治彦，粂直人，“秘密分散技術を用いた HIS バックアップクラウド環境の実現性評価,” 医療情報学, Vol.33, No.4, pp.255–233, 2013.

[6] 厚生労働省，“医療情報システムを安全に管理するために 「医療情報システム安全管理に関するガイドライン」 すべての医療機関等の管理者向け読本,” http://www.mhlw.go.jp/shingi/2009/03/dl/s0301-6b.pdf，2009.

[7] 健康保険組合連合会, “政策立案に資するレセプト文責に関する調査研究 (最終報告書),” http://www.kenporen.com/include/outline/pdf/chosa25_02.pdf, 2015.

[8] G. Blakley, “Safeguarding Cryptographic Keys,” Proc of AFIPS 1979 Nat. Computer Conf, Vol.48, pp.313-317, Sept. 1979.

[9] 山本博資, “ 秘密分散とそのバリエーション,” 数理解析研究講義録, 1361 巻, pp.19–31,2004.

[10] 山本博資，“(k, L, n) しきい値秘密分散システム,” 電子通信学会誌, Vol.J68-A No.9, pp.945–952, Sep. 1985.

[11] 尾形わかは，黒沢馨，“秘密分散共有法とその応用,” 電子通信学会誌, Vol.82, No.12,pp.1228–1236, Dec. 1999.

[12] 独立行政法人 情報処理推進機構 “アクセス制御に関するポリシーモデルの調査 報告書,” http://www.ipa.go.jp/files/000013734.pdf, 2005.

[13] 倉上高史, 藤吉正明, 貴家仁志,“秘密分散による段階的アクセス制御,” Technical Report ISEC2013-26, pp.193–198, 2013.