

薬歴データへのアクセスを想定した大規模災害時の 高可用ストレージ基盤の耐災害性評価

宗形 聡[†] 宋 チュウ[†] 手塚 大[†] 村岡 裕明[‡]

株式会社日立ソリューションズ東日本[†] 東北大学電気通信研究所[‡]

1. はじめに

東日本大震災では、地震や津波により建物が損壊し、沿岸部を中心に医療情報などの重要なデータが情報ストレージから喪失した[1]。震災直後の段階から慢性疾患患者へのケアが必要とされたが、医療機関からデータが失われた地域では、円滑な災害医療の提供が非常に困難であった。その一方で、過去の薬歴が記載されたお薬手帳を持つ患者に対しては、避難所にいる薬剤師等から迅速に薬の処方なされるなど、災害時に個人が薬歴を所持していることの有用性も明らかになった。お薬手帳は今後、国の IT 戦略の下で薬歴データを記録する電子お薬手帳としての普及が見込まれている。

電子お薬手帳は、個人の携帯端末がクライアントとなり、クラウドや医療機関等にサーバやストレージが配備される Web システムとして構築されることが多い。大規模災害時に被災者が自身の薬歴データを利用するためには、建物の被災によるストレージからのデータ喪失を防ぐこと、および被災地でインターネットが途絶しても代替手段によってストレージのデータを参照できること、以上の2点が Web システムには求められる。

こうした大規模災害に強い情報システムを実現するストレージとして、高可用ストレージ基盤（以下、ストレージ基盤）が開発されている[2]。ストレージ基盤は、限定された地域で互いに近隣に存在する複数の拠点ストレージ群から構成される。任意の2つの拠点ストレージはネットワークで相互接続されており、ある拠点ストレージに格納されたデータは、別の拠点ストレージに複製される。複製先の選択は、元データを保持する主拠点と複製データを保持する代替拠点が災害で同時に被災してデータを失わないように、2拠点間の同時被災リスクに基づいて実行される。ストレージ基盤は、災害で半数の拠点が損壊しても90%以上のデータにアクセスできる可用性を性能目標としている。

地域の医療機関等を拠点とするストレージ基盤上に電子お薬手帳システムを構築すれば、災害で多数の拠点が被災しても複製により喪失データを低減でき、インターネットが途絶してもデータを保持する近隣の拠点まで個人が直接移動して、拠点内 LAN 経由でストレージからデータを参照できると期待される。一方で、実際に構築された電子お薬手帳システムを用いて、大規模災害時に被災者が自身の薬歴データを問題なく参照可能か、実運用を想定した場面での耐災害性の評価が課題となっている。

本研究では、東日本大震災の被災地をフィールドとして、各拠点の被災状況やユーザによるシステムの利用状況を実態に合わせてモデル化し、当該モデルに基づきス

トレージ基盤の理論上の耐災害性を評価した。ストレージ基盤で実行されるリスクに基づく複製先選択と、既存の分散ストレージで採用されるランダムな複製先選択で、それぞれ耐災害性を比較した。その結果、リスクベースの選択によって耐災害性が向上し、ストレージ基盤上で大規模災害に強いシステムを実現できることが分かった。

2. 耐災害性評価方法の確立

評価に際して設定した運用状況は、仙台市および周辺自治体の一部からなる地域にある n 箇所の医療機関が拠点として活用され、当該地域で生活する m 人のユーザが電子お薬手帳の利用者となり拠点ストレージに薬歴データを格納している、というものである。

この設定の下でストレージ基盤の理論上の耐災害性を計算するため、実際に大規模災害が生じた場合の地域の被災状況と、ユーザのシステム利用状況をモデル化する。

2.1. フィールドの被災状況モデル

大規模災害が発生した直後のフィールドの被災状況として、まず、拠点集合 S に属する拠点 $s \in S$ の損壊確率 p_s を以下のように定式化した。

$$p_s = \sum_{c \in C} (e_s(c) \times d(c)) \dots (1)$$

ここで C は震度を表す集合で $C = \{\text{震度 4 以下, 5 弱, 5 強, 6 弱, 6 強以上}\}$ とする。 $e_s(c)$ は、当該地域で今後発生する全ての地震を含む地震動モデル[3]から求まる、拠点 s での c クラスの地震動確率を表し、 $d(c)$ は地震被害関数[4]により推定された c クラスの地震による建物全半壊率を表す。

また、拠点の損壊パターンを $\{0, 1\}^n$ のベクトルで表し、取りうる損壊パターンの集合を D とする。 D の要素数 $\#D$ は 2^n となる。 $\mathbf{x} = (x_1, \dots, x_n) \in D$ について、 $x_s = 1$ で拠点 s が損壊、 $x_s = 0$ で拠点 s が稼働を表す。このとき、ある損壊パターン \mathbf{x} の発生確率 $P(\mathbf{x})$ を、各拠点の損壊が独立事象であると仮定して、同時損壊確率で求める。

$$P(\mathbf{x}) = \prod_{s \in S} (p_s x_s (1 - p_s) (1 - x_s)) \dots (2)$$

次に、災害直後のインターネットの途絶状況として、災害後の経過時間 t におけるユーザのインターネット接続可能率 $I(t)$ を以下のようにモデル化した。

$$I(t) = \min\{PV(t)/\overline{PV}, 1\} \dots (3)$$

ここで、 $PV(t)$ は t 時点でのある Web サイトのページビュー、 \overline{PV} は災害前の同じサイトのページビューである。本研究では、Yahoo! Japan のビッグデータ分析レポートに記載された宮城県のスマートフォンの PV 値をもとに $I(t)$ を推計した。

2.2. ユーザのシステム利用状況モデル

大規模災害が発生した直後のユーザによる電子お薬手帳の利用状況として、災害後の経過時間 t におけるユーザの薬歴データ参照数 $R(t)$ を式 4 のようにモデル化した。

$$R(t) = r(t) \times R_T / T \dots (4)$$

ここで、 R_T は期間 1 から T までの m ユーザの参照総数であり、宮城県の月次の処方箋発行枚数や外来患者の平均通院間隔などの統計値から推計できる。 $r(t)$ は時間 t で

Disaster-tolerance Evaluation for a Highly-available Storage Platform Replicates Electronic Medication Record Data under Severe Conditions that Large Scale Earthquakes Occur

[†]Satoshi Munakata, Chong Song, and Masaru Tezuka, Hitachi Solutions East Japan, Ltd.

[‡]Hiroaki Muraoka, RIEC, Tohoku University

表 1. 実験の前提条件

拠点数 (ストレージ数)	$n=24$
電子お薬手帳ユーザ数	$m=100,000$
損壊拠点数	$k=12$
災害後の参照期間	$T=5$
インターネット接続可能率	$I(t)=\{0.52, 0.52, 0.79, 1.00, 1.00\}$
期間 1~ T までの参照総数	$R_T=33,200$
参照数の変動率	$r(t)=1.0$
参照ユーザ $u \in U(t)$ の決定	全ユーザから一様乱数で選択

の参照数の変動率を表し、 $\sum_t r(t) = T$ を満たす。 $r(t)$ は曜日や時間帯ごとの通院患者数の増減から推計できる。

2.3. 評価指標の算出

ストレージ基盤の耐災害性指標として、本研究では可用性を選択した。可用性の計算には NICT の分散ストレージシステムで定義された算出式[5]、

可用性 = システム稼働率 × ファイルアクセス成功率を参考にした。ストレージ基盤は n 個の拠点ストレージがネットワークを介して並列接続されているため、各拠点ストレージの稼働率を 0.99 と仮定しても、基盤全体の稼働率は $1 - (1 - 0.99)^n$ となり、 n が十分大きければ 1 と見なせる。そこで、ファイルアクセス成功率を定式化した。

災害後の経過時間 t ($t=1, \dots, T$) で電子お薬手帳を利用するユーザの集合を $U(t)$ 、ユーザ $u \in U(t)$ の薬歴データを保持する拠点の集合を $S_u \subset S$ とする。このとき、損壊パターン $\mathbf{x} \in D$ の下でユーザ集合 $U(t)$ によるデータ参照が成功する割合 $A(\mathbf{x})$ を以下の式 5 で定めた。

$$A(\mathbf{x}) = \sum_{t=1}^T \sum_{u \in U(t)} f(\sum_{s \in S_u} x_s) / \sum_{t=1}^T \#U(t) \dots (5)$$

ここで、 $f(z)$ は $z \geq 1$ の場合に 1、それ以外に 0 を返す関数である。また、 $\#U(t) = R(t)$ である。式 5 では、 u の薬歴データ (複製も含む) を保持する拠点の中で 1 つでも損壊していない拠点があれば、データを参照可能としている。その際、 u がインターネットに接続できる場合には、電子お薬手帳を用いてインターネット経由でデータを参照し、接続できない場合には、データを保持する拠点まで移動して拠点内 LAN 経由でデータを参照すると考える。

参照成功率 $A(\mathbf{x})$ はパターン発生確率 $P(\mathbf{x})$ の分布に従う確率変数と見なせる。そのため本研究では、ストレージ基盤の可用性目標にある半数の拠点損壊が発生する全ての損壊パターンを考慮し、期待される可用性 $E[A]$ を評価指標として定めた。一般に k 個 ($k \leq n$) の拠点が損壊するパターンの集合 D_k を $D_k = \{\mathbf{x} | \sum_{i=1}^n x_i = k\} \subseteq D$ とするとき、可用性 $E[A]$ は式 6 のように書ける。

$$E[A] = \sum_{\mathbf{x} \in D_{n/2}} (A(\mathbf{x})P(\mathbf{x}) / \sum_{\mathbf{x} \in D_{n/2}} P(\mathbf{x})) \dots (6)$$

3. 計算機実験

大規模災害直後のフィールドの被災モデル、ユーザによる薬歴データの参照モデル、および評価指標の算出モデルをプログラム実装し、表 1 に示す前提条件の下でストレージ基盤の理論上の可用性を計算した。評価に際し、ストレージ基盤のリスクベースの複製と、従来の分散ストレージで広く用いられているランダム複製でそれぞれ可用性を比較した。表 2 に両複製手法による複製先拠点の選択結果 (拠点番号) を示す。リスクベース複製では、式 1 に示す各拠点の損壊確率を用いて 2 拠点間の同時被災リスクを算出し、複製先を決定した。ランダム複製では、拠点ごとに一様乱数を用いて複製先を決定した。

表 2. リスクベース複製とランダム複製の複製先拠点

Site	Risk	Rnd	Site	Risk	Rnd	Site	Risk	Rnd
1	5	12	9	12	18	17	1	16
2	4	6	10	24	7	18	15	15
3	23	20	11	7	1	19	16	24
4	2	22	12	9	17	20	13	23
5	14	2	13	20	21	21	8	10
6	22	9	14	6	8	22	17	5
7	11	3	15	3	14	23	18	4
8	21	13	16	19	19	24	10	11

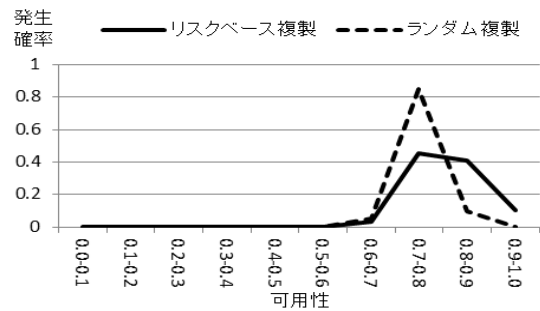


図 1. 半数拠点損壊時のストレージ基盤の可用性分布

全ての半数拠点損壊パターンに対して式 5 を計算して求めた理論上の可用性分布を図 1 に示す。図 1 からは、リスクベース複製を実行した場合のストレージ基盤の可用性は、ランダム複製を実行した場合よりも高くなる傾向にあることが見て取れる。式 6 に示す $E[A]$ についても、それぞれ 0.809 と 0.756 となり、 t 検定により 1% 水準で有意差が認められた。以上の結果から、リスクベース複製を実行するストレージ基盤により大規模災害に強い電子お薬手帳の実現が可能であることが分かった。

4. おわりに

本稿では、大規模災害が発生する地域について、災害後の被災状況やシステム利用状況をモデル化し、ストレージ基盤の理論上の耐災害性を評価した。評価指標には、半数拠点が損壊する全てのパターンから得られる可用性の期待値を用いた。計算機実験の結果、リスクに基づき複製先を決定するストレージ基盤は、ランダムに複製先を決定する場合よりも高い可用性を示した。ストレージ基盤上に電子お薬手帳を構築すれば、大規模災害が発生してもより多くの被災者が薬歴データを参照可能になる。

5. 謝辞

本研究の成果は、文科省委託事業「高機能高可用性情報ストレージ基盤技術の開発」の成果の一部である。

参考文献

- [1] 田中博, “災害時と震災後の医療 IT 体制 そのグランドデザイン”, 情報管理 Vol.5(12), pp.825-835, 2012.
- [2] S. Matsumoto, et al., “Risk-aware Data Replication to Massively Multi-sites against Widespread Disasters”, Proc. of the 2nd Asian Conference on Information Systems, 2013.
- [3] 防災科学研究所, “J-SHIS 地震ハザードステーション”, <http://www.j-shis.bosai.go.jp/>, 2014.11.19 accessed.
- [4] 翠川三郎, 他, “兵庫県南部地震以降の被害地震データに基づく建物被害関数の検討”, 日本地震工学会論文, Vol.11(4), 2011.
- [5] 渡邊英伸, 他, “NICT サイエンス〜クラウド広域分散ファイルシステムのセキュリティ機能拡張の要件〜”, 信学技報, Vol.113(86), pp.13-18, 2013.