

[2020]九州大学情報統括本部年報 : 2020年度

<https://hdl.handle.net/2324/4741344>

出版情報 : 九州大学情報統括本部年報. 2020, pp.1-, 2021-12-01. Information Infrastructure Initiative, Kyushu University

バージョン :

権利関係 :



第1章 応用データ科学研究部門

1.1 スタッフ一覧

職名	氏名	研究キーワード
教授	小野 謙二	数値流体力学、可視化、大規模並列計算、機械学習
准教授	鈴木 孝彦	演繹データベース、論理プログラミング、テキストマイニング、 数値データの異常検出
准教授	伊東 栄典	Webマイニング、Webサービス、情報検索、情報統合、XML、分散 システム、ネットワーク、協調システム、ソフトウェア工学
助教	内林 俊洋	クラウドコンピューティング、エッジコンピューティング、IoT、地域公 共交通、AIセキュリティ
准教授	櫻井 大督	可視化、機械学習、トポロジカルデータ解析

1.2 研究事例紹介

1.2.1 「低 B/F アーキテクチャに向けた連立一次方程式の高速解法の研究」

連立一次方程式の解法アルゴリズムは科学技術計算における基本的な研究対象の一つである。近年のシミュレーションの大規模化に伴い、巨大な連立一次方程式を扱うことが必要になってきたが、使用する計算機の性能向上に伴い、演算処理の性能に対してデータ移動性能が相対的に小さな、いわゆる低 B/F アーキテクチャの計算機での演算性能が低いという問題が生じている。特に、疎行列計算における反復法では、理論性能の 5% 以下の性能となることも珍しくない。大規模疎行列を係数行列にもつ連立一次方程式の計算性能を改善するため、低 B/F アーキテクチャに適した反復解法アルゴリズムを開発してきた。本研究では、非圧縮性流体解析において計算に時間がかかる圧力のポアソン方程式の反復解法部分を効率的に解く手法である Parallel Cyclic Reduction (PCR) 法について[ono:2020-pcr]報告する。PCR 法は、これまでマルチコアアーキテクチャである Intel Xeon Skylake-SP、「富岳」と同様なメニーコアアーキテクチャである SGI UV300、キャッシュ付きのベクトル型プロセッサである NEC SX-Aurora TSUBASA を用いたベンチマークを実施し、いずれのアーキテクチャにおいても高い実行性能を達成できるアルゴリズムであることを確認した。本年度は、PCR 法の更なる高速化、「富岳」でのチューニングと実行性能の確認を行った。

PCR 法は、三重対角行列の解法で最も効率的である LU 分解を元にした Thomas Algorithm を並列処理する解法である。Thomas Algorithm は逐次アルゴリズムであり、高速計算のために必須な SIMD 演算も利かない。その原因はループの依存性であり、PCR 法はその依存性を除去するために、Figure 1 に示すような「縮約」操作を繰り返し行う。1 回の「縮約」により、与えられた 1 組の N 元連立一次方程式は依存関係のない 2 組の $N/2$ 元連立一次方程式に分類される。この操作を繰り返すと、1 組 \rightarrow 2 組 \rightarrow 4 組 $\rightarrow \dots \rightarrow N$ 組の 1 個の方程式となる。最終段階では方程式は自明に解けることになる。

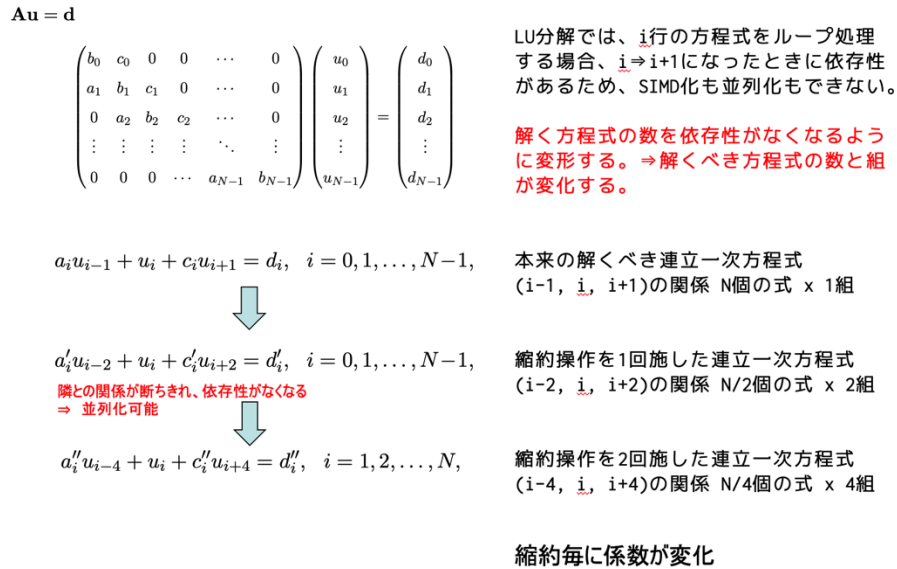


Figure 1 Reduction operation in Parallel Cyclic Reduction method.

Figure 2 に PCR 法のカーネルコードを示す。関数への引数 s は縮約の回数を示しており、縮約が進むと参照するメモリの距離が離れることがわかる。考えている連立一次方程式の数は Z 軸の格子点数なので、せいぜい 5 1 2 程度である。単精度の場合、 d, a, c, dn, an, cn の 6 配列を利用する実装となっているので、計算に必要な記憶領域は $4\text{Byte} \times 6 \times 512 = 12\text{kB}$ 程度である。テストした計算機の L1 キャッシュサイズは 32kB 程度あるため、計算に必要なデータはほぼ L1 キャッシュに載る。したがって、参照のスライドは大きくなるが、全て L1 キャッシュ内でアクセス可能となる。実際に Skylake-SP での実行時のプロファイル結果からは、Figure 3 のように PCR 法は他の古典的な方法に比べて圧倒的に L1 キャッシュのヒット率が高いことが確認でき、アルゴリズムの有効性の根拠を示している。

次に、PCR 法をさらに高速化するため、PCR の縮約回数を最終段から 2 段手前で止める実装を評価した。これは、縮約を繰り返すことによる精度低下を抑制することと、2 段手前で得られる 4×4 の方程式を直接反転することにより B/F をさらに低下させる狙いがある。Figure 4 にその効果を示す。ここではアウタースイープを昇順に走査する `pcr_esa`, Red-Black オーダリングで走査する `pcr_rb_esa` を比較した。Original と 4×4 を Cramer の方法で直接反転する実装を比較している。Cramer の方法はいずれもオリジナルよりも性能が改善されていることがわかる。

```

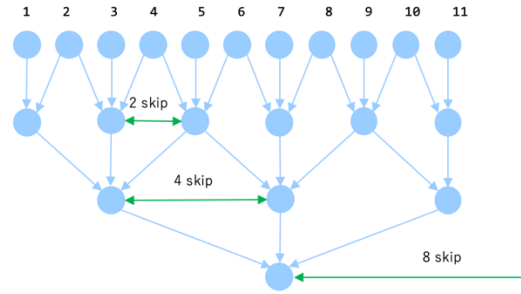
void CZ::pcr_kernel(const int nx, const int s,
                  REAL_TYPE* d, REAL_TYPE* a, REAL_TYPE* c,
                  REAL_TYPE* dn, REAL_TYPE* an, REAL_TYPE* cn,
                  double& flop)
{
    REAL_TYPE r, ap, cp;
    int iL, iR;

    flop += (double)(nx)*14.0;

    #pragma omp parallel for private(iL, iR, ap, cp, r)
    for (int i=1; i<=nx; i++)
    {
        iL = std::max(i-s,0);
        iR = std::min(i+s,nx+1);

        ap = a[i];
        cp = c[i];
        r = 1.0 / ( 1.0 - ap * c[iL] - cp * a[iR] );
        an[i] = - r * ap * a[iL];
        cn[i] = - r * cp * c[iR];
        dn[i] = r * ( d[i] - ap * d[iL] - cp * d[iR] );
    }
}

```



ステージにより $s=1, 2, 4, \dots$ と変化し、メモリアクセスの幅が広がる。最良の場合、 $c[i]$ と $c[iL]$ 、 $c[iR]$ は連続しているので1ロード、最悪の場合は3ロードとなる。

Figure 2 Pseudo code of PCR.

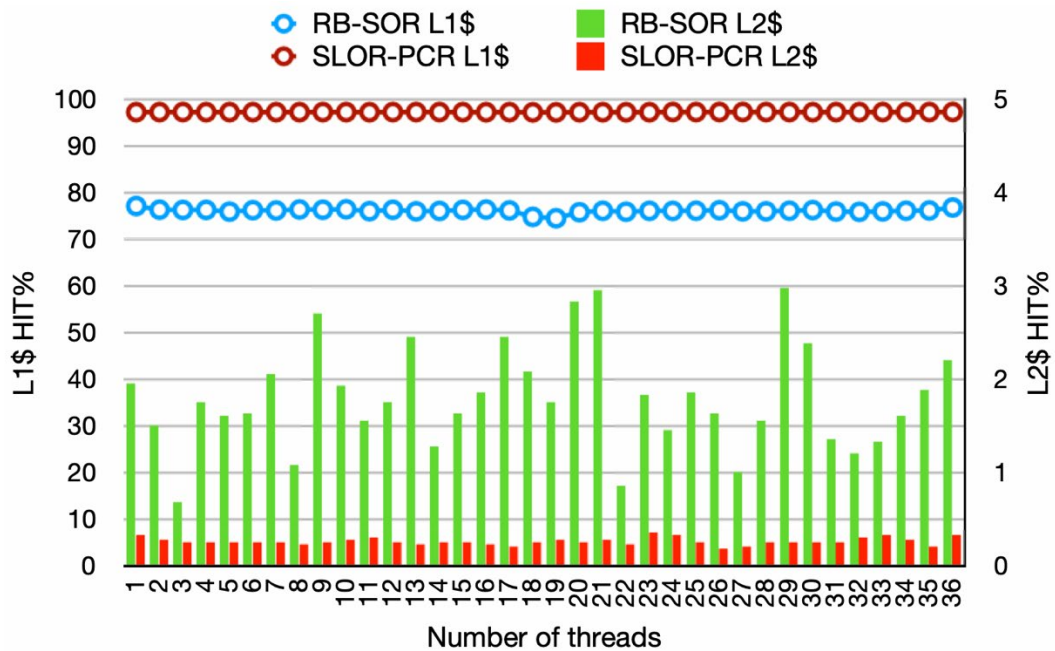


Figure 3 Cache hit rate for 36 cores on Skylake-SP that has 32kB L1 cache.

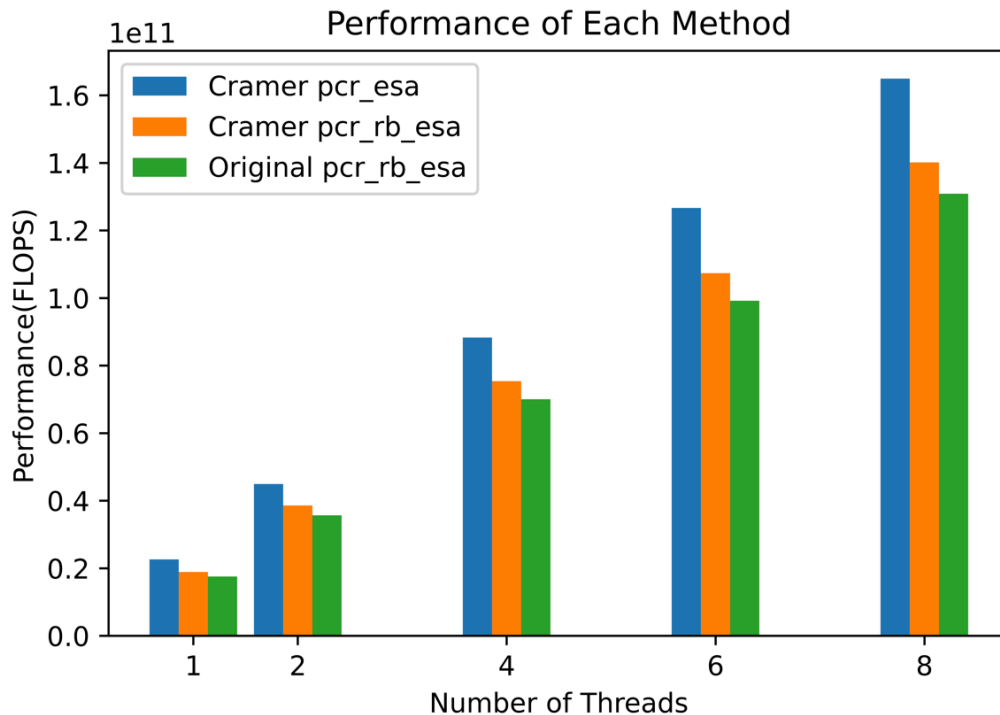


Figure 4 Performance improvement by direct inversion of 4x4 matrices.

次に、「富岳」でのプロファイリングに基づくチューニングを実施した。実施した内容は、不連続なメモリアクセスが生じる箇所でのプリフェッチを指示、ソフトウェアパイプライン最適化の2点である。測定結果を Figure 5 に示す。演算性能の測定は、「富岳」のプロファイラと PMlib による手動演算数カウントによる方法を載せている。プロファイラは投機実行などによる余計な演算器の動作も含んでいるため、プログラムループ中の演算数をカウントした測定よりも3割ほど多くなる点に注意する。1 CMG (12 スレッド) の場合も 4 CMG (48 スレッド) の場合もオリジナルよりも 1.23 倍以上の演算性能であることがわかる。メモリスループは 50 GB/s 程度でバンド幅を使い切っていないが、本アルゴリズムは L1 律速であるため、余裕がある。チューニング後の 48 コア実行時のハードウェアカウンタの値は推定 1 TFLOIPS となり、ピーク性能 29% となる。

以上のように、提案する PCR 法は L1D キャッシュ負荷が大きく、低 B/F、動作周波数の高い CPU 向け、キャッシュを有効利用できる、メニーコアの並列性を引き出せるアルゴリズムであり、「富岳」においても高い性能を発揮することが確認できた。

■ 計算性能

機種	ソースコード	プロセス構成 [スロット×プロセス]	SIMD 命令率	演算性能※1 [Gflops]	メモリスループット [GB/秒]	実行時間※2 [秒]
FX1000	リジナル	12×1	61.10%	220.03 (165.63)	53.15	1.784 (1.00)
		48×1	—	(616.85)	—	0.479 (3.72)
	チューニング あり	12×1	60.23%	272.89 (205.31)	50.87	1.440 (1.24)
		48×1	—	(792.77)	—	0.373 (4.78)

- 演算性能の括弧内はPMlibの値(手動カウント)
- ハードウェアカウンタ(HWC)は3割+多い値を示す
- FX1000の48コア実行時のHWC値は推定1,000Gflops、ピーク比29%の実行性能
- PCRソルバーはL1Dキャッシュ負荷が大きく、低B/F、動作周波数の高いCPU向き
- PREFETCHが効く実装

Figure 5 Performance results on an A64FX.

[ono:2020-pcr] Kenji Ono, Toshihiro Kato, Satoshi Ohshima, Takeshi Nanri, “Scalable Direct-Iterative Hybrid Solver for Sparse Matrices on Multi-Core and Vector Architectures,” HPCAsia2020: Proceedings of the International Conference on High Performance Computing in Asia-Pacific Region, pp.11-21, <https://doi.org/10.1145/3368474.3368484>, 2020.

1.2.2 「Benford の法則 の実数底への拡張の試み」

1. Benford の法則の応用事例

2020 年末の米大統領選挙では、票数が不正に操作されているのではないか？という話題が大きく取り上げられた。不正検出のツールとして、会計監査で使われている Benford の法則が一部で使われた[1]。

Benford の法則は自然な数値の集合においては、最上位桁 (First Significant Digit(FSD)) の分布が一様でないことに基づいている。たとえば、市町村人口統計は Benford の法則に従い、最上位桁 1 の数値 (ex. 1, 587, 167, 252, …) の出現確率 p は $\log_{10} 2 = 0.301$ に非常に近いのに対し、最上位桁 9 の数値 (ex. 912, 9, 600, 93, 456, …) は、 $p = 1 - \log_{10} 9 = 0.046$ に近い。

米大統領選の得票分布を調査したところ、一部で Benford の法則からの乖離が発見された。しかしながら、これが票数の不正な操作によるものなのか、他の原因によるものなのかはおおいに議論の余地がある。

過去の選挙においても、票数の分布は Benford の法則から乖離している場合があることが指摘されており、得票数分布特有の性質に基づく乖離ではないかとも考えられる。

ある数値集合の分布が Benford の法則から乖離している場合、その要因を具体的に特定できれば、不正検出だけでなく、その数値集合特有の隠れた性質の発見にも役立つ可能性がある。

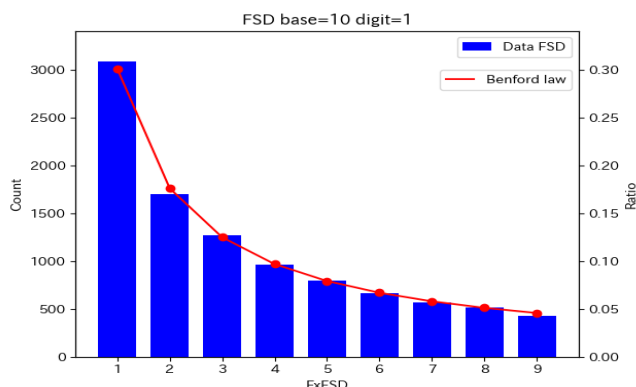


図1: Benford の法則に従う数値データ
最上位桁の出現ヒストグラム

2. Benford の法則の拡張 (k 進数、上位 n 桁)

Benford の法則は数値データの対数表現と強く関係している[2]。10 進数は自然対数表現に対応し、対数の底 k を 9, 8, 7, …, 11, 12, 13 と変化させればそれぞれの k において Benford の法則を導くことができる。また、Benford の法則は、最上位桁についてだけではなく、上位 $n=2, 3, \dots$ 桁についても成り立つことが知られている。この 2 つを組み合わせると、例えば、2 進数表現の上位 4 桁についての Benford の法則を考えることができる。

この拡張には 2 つのメリットがある。第 1 に、複数の (k, n) の組について、Benford の法則からの乖離を調べれば、データのどの部分が乖離の要因になっているかを特定することができる。第 2 に、従来 10 進数の Benford の法則が適用できなかった数値の集合についても、Benford の法則からの乖離を調べることができる。

図 2 に、 $\text{Avr.} = 50$, $\text{S.D.} = 18$ の正規分布に従うように発生させた 10,000 個の乱数データのヒストグラムと、数値を 2 進数表現したときの上位 4 桁の出現回数・確率を示す

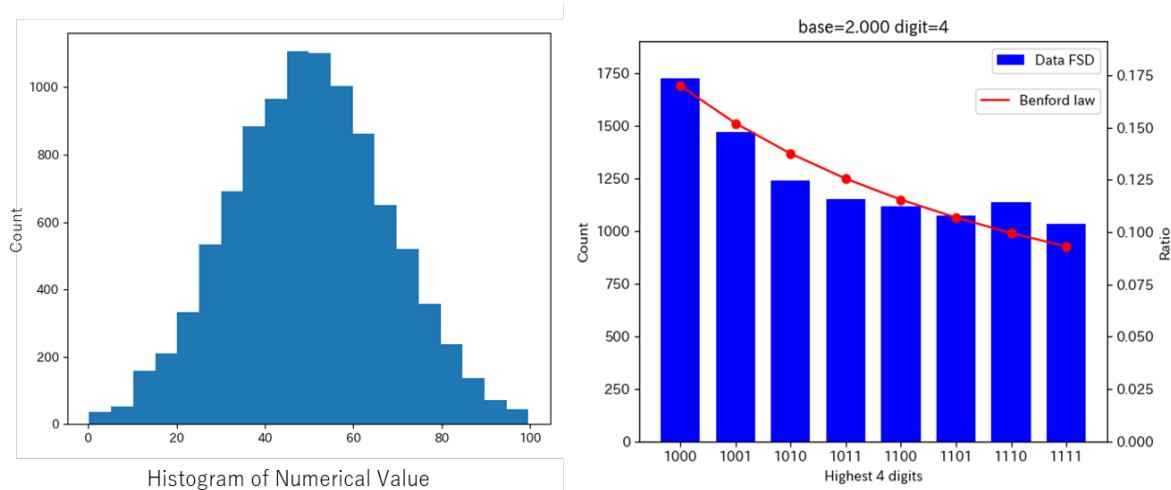


図 2: Avr. =50, S. D. =18 の正規分布乱数データと、データの 2 進数表現上位 4 桁の出現ヒストグラム

図 2 で示した数値集合では、40~60 のピーク に多くの数値要素が含まれるため、10 進数での Benford の法則が成り立たないのは自明である。ところが、2 進数上位 4 桁については、Benford の法則に近い部分と、法則から乖離した部分が存在する。複数の底 k について乖離を調べることができれば、乖離の要因となる部分を特定できる可能性がある。

しかしながら、このデータでは 3 以上の底については、Benford の法則からの乖離が大きくなり、複数の整数底 k を組み合わせた分析は困難である。

3. 実数底への拡張

Benford の法則は、次のように定式化することができる

x を k 進法で表わしたときの最上位 n 桁 $FSD(x, k, n)$

$$FSD(x, k, n) = \text{floor}(\text{pow}(k, \text{frac}(\log(x, k)) + k - 1))$$

Benford の法則に従う FSD $m (= 1, 2, \dots, k-1)$ の出現確率 $P(m, k, n)$

$$P(m, k, n) = \log(m + 1, k) - \log(m, k)$$

ここで、 $\text{pow}(k, \text{frac}(\log(x, k)) + k - 1)$ は、 x を k 進数の仮数+指数表現にした時の仮数部 (significant) にあたる。 $k-1$ は仮数部を $[k^{n-1}..k^n)$ に正規化するための係数である。floor 関数は、 k 進数仮数の上位 n 桁を取り出し、区間を等分割するために使われている。

以上から、 k 進数における Benford の法則を言い換えると

「数値 x を k 進数の仮数*指数表現 $x = y * \text{pow}(k, z)$ ($y = [k^{n-1}..k^n)$) で表現した場合、確率

$$\Pr(y1 \leq y < y2) = \log(y2, k) - \log(y1, k)$$

である」となる。この性質は、 k が整数以外でも成り立つと予想できる [2]。

図 3 に、図 2 と同じデータについて、 $k = 2, 1, \dots, 2.4$ について、仮数部の出現ヒストグラムと Benford の法則から予測される値を示す。

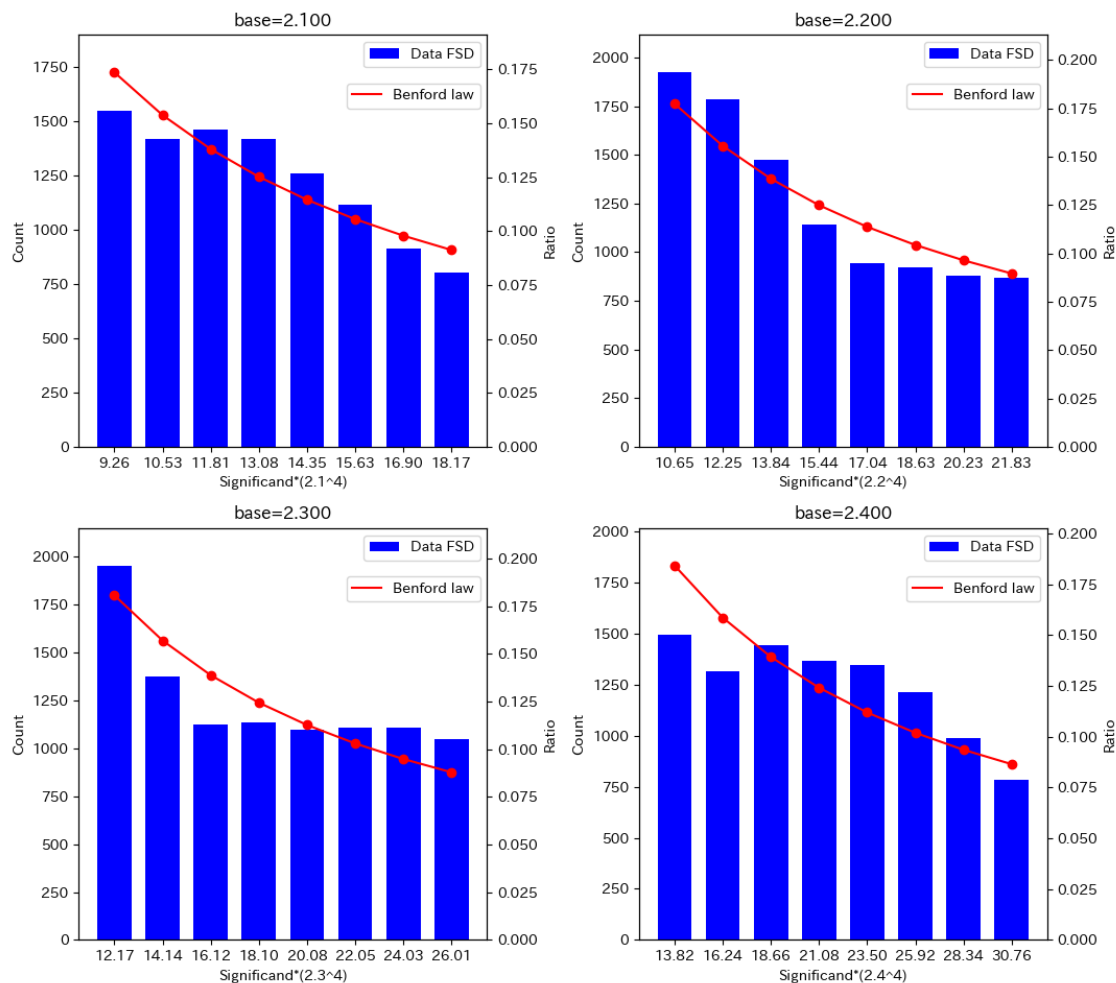


図3：正規分布乱数データの仮数部（底 $k = 2.1, 2.2, 2.3, 2.4$ ）ヒストグラムと Benford の法則への適合

複数の底を組み合わせることによって、元の数値集合のどの部分が Benford の法則からの乖離に影響しているのか調べる事が可能になる。現在、従来 Benford の法則の適用範囲外とされてきた試験の点数などの実データについて、調査を行っている。

[1] Fact check: Deviation from Benford’s Law does not prove election fraud

<https://www.reuters.com/article/uk-factcheck-benford-idUSKBN27Q3AI>, 2021/5/24 閲覧

[2] A Short Introduction to the Mathematical Theory of Benford’s Law, Berger, A and Hill T.P.

In Benford’s Law Theory and Applications ed. Miller S. J., Princeton University Press, ISBN:978-691-17461-1, 2015

1.2.3 「ブロックチェーンを利用した VM マイグレーション制御機構の研究」

IoT が普及したことで、個人に紐づいた環境を測定する各種センサや個人の健康状態を測定するセンサからの膨大な量のデータをクラウドへ保管することが一般的となった。IoT で収集されたこれらの膨大な量のデータは、個々の値だけでなく、複数のデータを組み合わせることで、個人の嗜好や行動の予測、行動への介入等に有用な手段を明らかにするため解析がサービス提供者により行われている。これらのデータは個人に密接なデータであり、管理に細心の注意が必要となる。データの所有者の許諾した条件だけでなく、GDPR を代表とする国や地域の法規制、組織の規約を遵守した適切な管理と利用が不可欠である。また、IoT だけでなく、ネットワークを介するアプリケーションやサービスを運用するにあたり、クラウド環境を利用したサービスが一般的になっており、クラウド環境なしでの運用は考えられないほど世の中に浸透している。その中でも、IaaS クラウドの利用時に特に注意が必要なことは不適切なデータの移動である。現在のクラウドにおいて、物理ホストマシン間を移動する VM マイグレーションは欠かせない。クラウドを提供するプロバイダにおいては、メンテナンスなどで物理ホストマシンを停止する際に、一時的に VM を他の物理ホストマシンへ退避させるなどの用途でマイグレーションを利用する。クラウドの利用者においては、エンドユーザに地理的に近い場所へ VM を移動させる際になどに利用する。これらの VM マイグレーションは非常に簡単に実行することができ、VM の管理者の契約の範囲で制限を受けるのみである。そこで問題となるのが、VM 内のデータやアプリケーションの許諾である。

利用時、特に注意が必要なことは無意識に行われる不適切なデータの移動である。クラウド環境内にあるホストマシン上に仮想マシンはデプロイされ稼働している。ホストマシン間を仮想マシンがマイグレーションすることは容易であり、仮想マシンの管理者は、必要に応じて仮想マシンが稼働するホストマシンを変更することができる。ホストマシンがすべて同一の場所にある場合は問題がないが、ホストマシンが複数国に渡って設置されている場合は法規制や組織の規約の違反を引き起こす可能性がある。しかし、VM マイグレーション時に VM 内のデータの許諾に関するチェックを行う仕組みは存在せず、不適切なデータの移動が発生する懸念がある。これは悪意のあるデータの移動だけ関わらず、無意識に不適切なデータの移動を行う危険性を含んでいる。そこで、VM マイグレーション時に、データの所有者の許諾した条件や国の法規制・組織の規約を確実に遵守するための仕組みを提案する。提案した仕組みをブロックチェーン内に構築することで、悪意のある改ざんを防ぎ堅牢な VM マイグレーション制御を可能とする(図 1)。

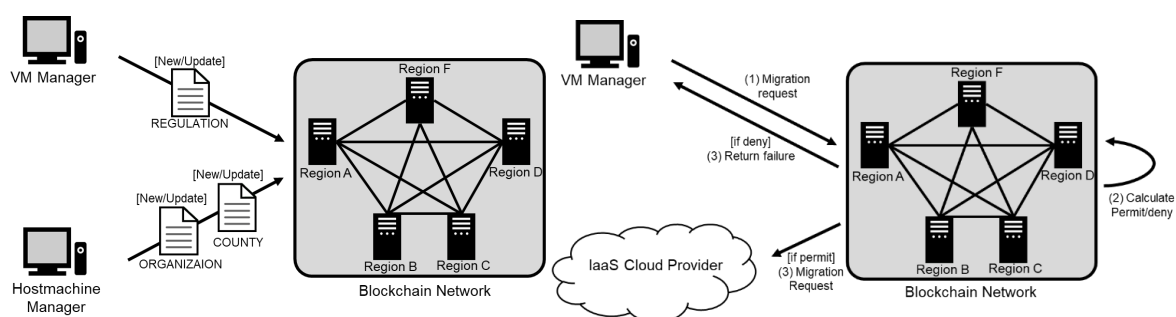


図 1 ブロックチェーンを利用したポリシー管理とマイグレーションの可否判定

提案したブロックチェーンを使用した制御機構を実装する。本機構は、マイグレーションのプロセスを変更することなく、さまざまな種類のクラウド環境に適用することができる。実装環境の構成を図2に示す。データ保護機構は、「VM 管理者」、「フロントエンド」、「クラウドプロバイダ」で構成する。「クラウドプロバイダ」は3つのリージョンを持つ5台のホストマシンで構成される。HostMachine01とHostMachine03とHostMachine04にはBCM(H(Block-chain Management Host))が設置され、ブロックチェーンにおけるノードの役割を兼任する。また、HostMachine02にはVM1が稼働している。「フロントエンド」は、クラウド環境内のブロックチェーンへの操作とクラウド環境への操作を行う。マイグレーション実行要求は「VM 管理者」が「クラウドプロバイダ」へ直接行うのではなく、フロントエンドへマイグレーション要求を行う。「フロントエンド」はブロックチェーンクライアントとクラウドマネージメントクライアントで構成される。ブロックチェーンクライアントは、クラウドプロバイダ内のBCM(H)とブロックチェーンネットワークを形成している。クラウドマネージメントクライアントは、クラウドプロバイダのSDKを操作でき、クラウドプロバイダへマイグレーションの要求を行う。具体的には、OpenStack APIを操作するために使用する。

また、各ホストマシン及びVM1の構成を表に示す。クラウドプロバイダはOpenStack Tikaで構築済みのもを使用する。ブロックチェーンはHyperledger Fabric 1.2とHyperledger Composer 0.20.1を使用する。ホストマシンとVM1のOSはCentOS7を使用する。また、ホストマシンは4コアCPU、12GBメモリ、100GBストレージで構成する。VM1は、1コア仮想CPU、2GB仮想メモリ、20GB仮想ストレージで構成する。

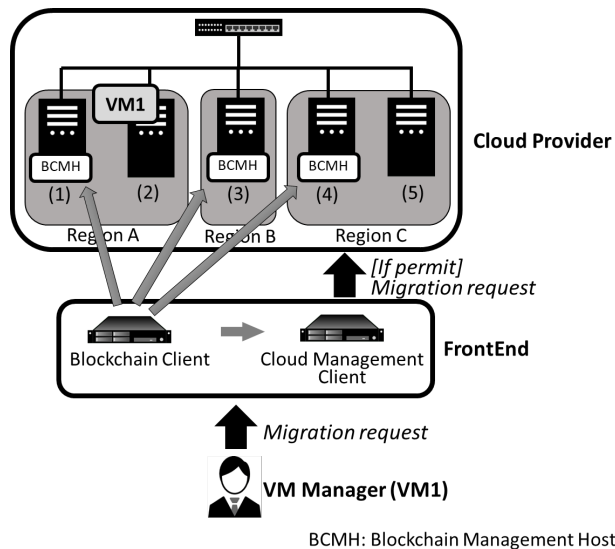


図2 提案した制御機構の実装

VMマイグレーション時に、データの所有者の許諾した条件や国の法規制・組織の規約を確実に遵守するための仕組みを提案した。提案した仕組みをブロックチェーン内に構築することで、無意識による不適切なデータの移動や悪意のある改ざんを防ぎ堅牢なVMマイグレーション制御を可能とする。提案した仕組みを既存のクラウド環境へブロックチェーンとともに実装した。VMとホストマシンに関するポリシーをブロックチェーンに登録し、ポリシーに準拠したVMライブマイグレーションの可否の判断が正確に行われているかを評価した。また、提案した仕組みを利用したライブマイグレーションの実行時間を計測し、オーバヘッドコストがないことを確認した。

1.2.4 「ファイバーのトポロジーとデータ解析」

計算機シミュレーションやデータ計測を行うと、物理的な現象の空間構造を解析する機会が多い。例えば気象では、雲が発達していく様子は、非常に微視的な空間スケールや巨視的なスケールではある程度わかっているのだが、その間のスケールで何が起きているのかはわかっておらず、雲の空間構造はひとつの手がかりとなる。そこで、データから空間構造を抽出していく計算の需要がある。このような空間構造の計算は各科学分野で発達していくのであるが、計算機科学では、特に、可視化研究に端を発することが多い。というのも、可視化の今ある研究分野は、90年代に、データ解析の合理化のためにコンピュータグラフィクスを採り入れたことから始まっており、ジェネリックな空間構造解析を研究することでデータ解析を促進しようとしてきた歴史があるからである。他分野で発達したアプローチと比べると、数値・数理的に厳密な手法やアルゴリズムを開発しているという特色がある。

しかし、データの空間構造を計算するのはなかなか難しい作業である。特に計算機科学の立場で問題を考えると、まず、第一に、解析結果から主観性を排除しにくい問題がある。第二に、異なる物理場のデータの相関をどういうやり方で解析し理解すべきかも、難しい。

まず、第一の問題の解決策の一つとして、トポロジーベースのデータ解析がやはり90年代から可視化に導入され、発達してきた。物理場、すなわち空間に分布する数値のデータについては、しきい値を決めて、ちょうどその領域に対応する領域を抽出する、コンタ解析が多い。このとき、抽出される曲面（しきい値の等値面）は、解析者が指定するしきい値の選択に主観性が入る。そこで、しきい値の選択に応じてどのようなコンタが抽出されるかひと目で把握することが重要視された。そのひとつの手段として、各しきい値での等値面の数（厳密には連結成分の数）がいくつかなどの情報を計算するのに、等値面の接続性（トポロジー）が計算された。また、しきい値を時間のように見ると、ある「時刻」で等値面が生まれ、それがその後の時刻で消えたりする。あるいは、ひとつの等値面が2つに分岐したり、逆に2つの等値面がひとつに合併したりする。こういうトポロジーの情報は、数学的にはReebグラフで表現できる。その後徐々に、可視化でこのようなトポロジーベースのデータ解析が発達していき、アルゴリズムが効率化したり、ノイズが除去されたり、あるいは新たな情報を付加したりなどの展開をしていっている。このような可視化コミュニティでの発達と並行して、計算幾何でも、点群から幾何構造を抽出する、「トポロジカルデータ解析」と呼ばれるデータ解析が、見事に成長していった。数理的には共通部分がかなり多いため、可視化コミュニティとのトポロジー計算との垣根は徐々に取り払われており、可視化研究の当事者はあまり区別にこだわらなくなっている。

さて、第二の問題について、トポロジカルなデータ解析はどう対応していくのであろうか。伝統的に、Reebグラフを用いる解析アルゴリズムは、残念ながら物理場をひとつと仮定しており、他の物理場を考慮することが非常に難しい。そこで、まず数理的に、Reebグラフの概念を一般化し、次に専用のアルゴリズムや解析手段を研究するアプローチが取られた。Reebグラフが、1つの場の等値面を表現するのに対して、2つの場の等値面の交差を表現するReeb space [1] が提案された。ここで、等値面の交差は、空間に分布しており、それぞれの場の等値面を定めるしきい値をそれぞれ1つずつ決めると定まる。Reeb spaceを使うと、さまざまなしきい値で定まる交差同士が空間の中でどう接続しているかを調べることができる。これは、2つの場にいる等値面の相互位置を解析する手がかりとして、データ解析AIの改善のために、当センターで研究が行われている [2, 3, 4]。

また、Reeb space の計算アルゴリズムは、これまでいくつか発表されてはいるが、まだ計算量が高い他、数理的な性質も、特に計算論的な視点から重要な未解明部分が残っている。そこで、当センターでは、新しい計算アルゴリズムや数理理論の研究も行っている。また、このような基礎的な研究はもちろん重要ではあるが、応用の裾野を広げることも同様に重要である。そこで、数理・計算・応用の専門家や企業からキープレイヤーを集め、マス・フォア・インダストリ研究所と共同で研究集会「Fiber Topology Meets Applications」[5]を開催するなどの活動も行っている他、国際連携でソフトウェア Toplogy ToolKit の研究開発及び提供 [6]も行っている。

参考文献

1. Edelsbrunner, H., Harer, J., and Patel, A. K. Reeb spaces of piecewise linear mappings. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*(pp. 242-250), 2008.
2. 櫻井大督. 文部科学省科研費若手研究 20K19809 (令和2~4年度) 仮説のオントロジーに基づく可視化。
3. Sakurai, D., Ono, K., Carr, H., Nonaka, J., and Kawanabe, T.: Flexible Fiber Surface: A Reeb-Free Approach. *Topology-Based Methods in Data Analysis and Visualization V*, Springer, pp. 187-201, 2020.
4. Sakurai, D., and Yamamoto, T., Visually Evaluating Topological Equivalence of Bounded Bivariate Fields, *Topological Methods in Visualization - Applications and Software*, in print.
5. 令和2年度九州大学マス・フォア・インダストリ研究所「産業数学の先進的・基礎的共同研究拠点」一般研究-研究集会(II)(令和2年度) Fiber Topology Meets Applications
6. Masood, T.B., Budin, J., Falk, M., Favelier, G., Garth, C., Gueunet, C., Guillou, P., Hofmann, L., Hristov, P., Kamakshidasan, A., Kappe, C., Klacansky, P., Laurin, P., Levine, J.A., Lukaszcyk, J., Sakurai, D., Soler, M., Steneteg, P., Tierny, J., Usher, W., Vidal, J., Wozniak, M., "An Overview of the Topology ToolKit," *Topological Methods in Visualization - Applications and Software*, in print.

1.3 研究内容紹介

1.3.1 小野 謙二

研究内容

- ・ 研究 — 数値流体力学、可視化、並列計算、遺伝的プログラミング
- ・ 教育 — 数値解析および演習、並列アルゴリズム、高性能並列計算法特論、オートモーティブ人間科学概論
- ・ 業務 — スーパーコンピュータの運用、サポート
- ・ ほか — 文部科学省科学技術試験研究委託事業、CREST、科研費、JHPCN、HPCI コンソーシアム理事

所属学会名

情報処理学会，日本計算工学会，可視化情報学会，日本機械学会，自動車技術会，日本流体力学会，IEEE，ACM

主な研究テーマ

- ・ In-situ / In-transit 可視化/データ処理基盤の研究開発
キーワード：可視化システム，並列処理，ユーザ利便性，リモート処理，2018.04～2023.03.
- ・ 深層学習を用いたデータからの自然法則の発見
キーワード：深層学習，遺伝的プログラミング，Lasso，2017.04～2020.10.
- ・ 機械学習による乱流解析手法の再構築
キーワード：深層学習，LES 乱流モデル，数値流体力学，2018.10～2020.10.
- ・ 時間並列計算法の研究
キーワード：時間方向マルチグリッド，Parareal 法，2015.10～2022.3.
- ・ シミュレーション実行支援環境の構築
キーワード：ワークフロー，データ管理，マルチプラットフォーム，エコシステム，2012.10～2020.10.
- ・ 複雑形状周りの熱流体流れシミュレータの開発
キーワード：直交格子，格子生成，1996.04～2026.12.

研究業績

● 原著論文

1. Xin Liang, Hanqi Guo, Sheng Di, Franck Cappello, Mukund Raj, Chunhui Liu, Kenji Ono, Zizhong Chen, Tom Peterka, Toward Feature-Preserving 2D and 3D Vector Field Compression, 13th IEEE Pacific Visualization Symposium, PacificVis 2020
2020 IEEE Pacific Visualization Symposium, PacificVis 2020 – Proceedings,
10.1109/PacificVis48177.2020.6431, 81-90, 2020.06.
2. Tomohiro Kawanabe, Kazuma Hatta, Kenji Ono, ChOWDER – A New Approach for Viewing 3D Web GIS on Ultra-High-Resolution Scalable Display, 2020 IEEE International Conference on Cluster Computing, CLUSTER 2020, 412-413, 9229604, 2020.09.
3. 小野謙二, 古賀壺成, 遺伝的プログラミングによる支配方程式の推定, Transactions of JSCES, <https://doi.org/10.11421/jsces.2020.20201004>, 2020.11.

● 学会発表

1. Xin Liang, Hanqi Guo, Sheng Di, Franck Cappello, Mukund Raj, Chunhui Liu, Kenji Ono, Zizhong Chen, Tom Peterka, Toward Feature-Preserving 2D and 3D Vector Field Compression, 13th IEEE Pacific Visualization Symposium, PacificVis 2020, 2020.06.
2. Tomohiro Kawanabe, Kazuma Hatta, Kenji Ono, ChOWDER A New Approach for Viewing 3D Web GIS on Ultra-High-Resolution Scalable Display, 2020 IEEE International Conference on Cluster Computing, CLUSTER 2020, 2020.09.

研究資金

● 科学研究費補助金

1. 2018年度～2020年度, 静岡県産業振興財団 次世代自動車技術革新対応促進助成事業, 分担, 燃料電池車用サイクロンセパレータの粉粒体分離効率の向上を目指した内部流れ解析及び最適化
2. 2016年度～2020年度, 厚生労働科学研究費補助金 (厚生労働省), 分担, 動的過程を扱う連成計算機構の研究開発と流体音数値解析への応用

● 学内資金・基金等

1. 2020年度～2021年度, 概算要求 教育研究活動 (取組), 代表, データサイエンスと異分野融合によるマルチエキスパート人材育成事業
ー汎オミクス計測・計算科学拠点の組織整備ー
2. 2019年度～2020年度, 2019年度大学改革活性化制度, 代表, 汎オミクス計測・計算科学アプローチに基づく異分野融合研究推進

教育活動

● 担当授業科目

1. 2020 年度・後期, 情報科学講究
2. 2020 年度・前期, 数値解析
3. 2020 年度・前期, 数値解析演習
4. 2020 年度・後期, 並列アルゴリズム
5. 2020 年度・前期, 【修士】高性能並列計算法特論
6. 2020 年度・前期, 情報学読解
7. 2020 年度・前期, 情報学論述 I
8. 2020 年度・前期, 情報学論議 I
9. 2020 年度・前期, 【博士】高性能並列計算法特論
10. 2020 年度・後期, 情報学演示
11. 2020 年度・後期, 情報学論述 II
12. 2020 年度・後期, 情報学論議 II
13. 2020 年度・通年, 国際演示技法
14. 2020 年度・通年, 知的財産技法
15. 2020 年度・通年, ティーチング演習
16. 2020 年度・通年, 先端プロジェクト管理技法
17. 2020 年度・通年, Scientific English Presentation
18. 2020 年度・通年, Intellectual Property Management
19. 2020 年度・通年, Exercise in Teaching
20. 2020 年度・通年, Advanced Project Management Technique
21. 2020 年度・通年, 基礎情報学特別講究
22. 2020 年度・通年, Advanced Research in Foundations of Informatics
23. 2020 年度・通年, 情報学特別講究第一
24. 2020 年度・通年, 情報学特別講究第二
25. 2020 年度・通年, 情報学特別演習
26. 2020 年度・通年, Advanced Research in Informatics I

第1章 応用データ科学研究部門

27. 2020年度・通年, Advanced Research in Informatics II
28. 2020年度・通年, Advanced Seminar in Informatics
29. 2020年度・前期, オートモーティブ人間科学概論

大学運営

- 学内運営に関わる各種委員・役職等

1. 2019.04～2021.03, 情報基盤研究開発センター附属汎オミクス計測・計算科学センター長
2. 2020.04～2022.03, 情報基盤研究開発センター長

1.3.2 鈴木 孝彦

研究内容

- ・ 九州大学教務事務システムの作成と運用支援
- ・ 九州大学学務情報システムの計画支援
- ・ 日本語 WordNet およびテキストマイニング
- ・ 機械学習
- ・ 異常検知

所属学会名

情報処理学会

研究業績

● 原著論文

1. Sachio Hirokawa, Takahiko Suzuki, Tetsuya Nakatoh, Prediction of Category of Scientific Article by Graph Convolution, Proc . ESKM 2020, 2020.09.

教育活動

● 教育活動概要

1. プログラミング言語特論 システム情報科学研究院
2. 情報処理概論 工学部エネルギー科学科
3. 情報処理概論 工学部物質科学科(材料コース)

● 担当授業科目

1. 2020 年度・後期, Advanced Programming Languages
2. 2020 年度・後期, 【博士】プログラミング言語特論
3. 2020 年度・後期, 【修士】プログラミング言語特論
4. 2020 年度・後期, 情報処理概論

1.3.3 伊東 栄典

研究内容

- ・ 情報検索・情報統合・情報連携

Web 上や電子データとして蓄積されている膨大なデータから、意味のある知識抽出や、抽出した知識を統合する研究を行っている。具体的には、Web データからの情報抽出、利用者コメントからの知識発見、コメントやリンク構造を利用したコンテンツ推薦などを行っている。ソーシャルブックマークからの新規情報の発見や、「クチコミ」と呼ばれる情報サービスからの知識抽出について研究している。

- ・ 集合知を利用した高品質コンテンツ検索

情報検索および知識発見に関する研究の具体的な対象として、ネット上のコンテンツを対象とした研究を行なっている。近年、ユーザ投稿型のコンテンツサービスが普及しており、動画・写真・静止画・小説などが増大している。これらのサービスでは、投稿機能だけでなく、視聴者からのコメント・タグ付け・リンクなどのフィードバック機能もある。視聴者からのフィードバックは、Folksonomy や集合知と呼ばれるもので、これを活用したコンテンツ検索を行なっている。

- ・ 電子認証基盤および認証フェデレーション構築

特定メンバーへの情報提供サービスや、オンラインでの商取引および申請のように、利用者を正しく認証する電子認証が求められている。本研究では電子認証機構および認可機構の構築と、情報サービス連携のための認証連携機構の構築を目的としている。多様な組織が柔軟にサービスを提供しあうための、電子認証基盤の構築についての研究開発を行う。その実現のために、同じポリシー・標準規格での認証基盤となるフェデレーションについての研究開発を行う。電子認証証基盤として、認証のためのデータベース構築、アプリケーションとなる情報サービスでの認証・認可機構の実現などが課題となる。組織内で使う、柔軟な利用者認証・認可システムの実現を目指す。

所属学会名

情報処理学会，電子情報通信学会

主な研究テーマ

- ・ 大規模データ解析

キーワード：大規模データ，分散処理，統計解析，クラウド・コンピューティング，
2012.06～.

ネットコンテンツの傾向分析

キーワード：ネット，コンテンツ，動画，小説，つぶやき，傾向分析，感情分析
2018.04～.

- ・ 情報検索・情報統合・情報連携

キーワード：情報検索，情報統合，Webマイニング，情報抽出，推薦，2000.04～.

教育活動

● 担当授業科目

1. 2020年度・夏学期，プログラミング演習 I (A)
2. 2020年度・前期，情報知能工学演習第二
3. 2020年度・前期，情報知能工学講究第二
4. 2020年度・後期，人工知能

大学運営

● 学内運営に関わる各種委員・役職等

1. 2012.06～，情報統括本部・認証基盤事業室
2. 2011.04～，情報統括本部・全学基本メール事業室

1.3.4 内林 俊洋

研究内容

1. 5G時代のエッジコンピューティング環境における情報保護制御機構の開発

エッジの仮想マシンやコンテナで稼働するアプリケーションが内包する利用許諾条件や組織の規約、国の法規制を、意識せずに遵守するための制御機構の開発を目指しています。

2. 仮想マシンマイグレーションにおける情報保護制御機構の開発

仮想マシンのマイグレーションを対象としたデータ保護機構を開発しています。具体的な環境構築に、クラウド基盤のOpenStackやブロックチェーンのHyperledger Fabricを使用して、セキュアな基盤の構築を目指しています。

3. マルウェア検知システムへのポイズニング攻撃とその対策

現在のAIを使ったマルウェア検知システムは、学習時に悪意のあるデータ(毒データ)が混入していることを想定しない。そこで、毒データを混入するための攻撃手法やその対策について研究しています。

4. 地域公共交通の支援活動

地域公共交通は基本的に人材不足や財源不足に悩んでいる。そこで、ICTを使った支援を行うことで、これらの問題を解決しようとしています。具体的には、コミュニティバスのロケーションシステム、乗降人数カウントアプリ、デジタルサイネージ、バス停ごとの乗降客数分析、そして従業員支援アプリケーションの開発など多岐にわたって支援を行っています。

所属学会名

日本情報経営学会, 電子情報通信学会, 情報処理学会, IEEE

主な研究テーマ

- ・ 5G時代のエッジコンピューティング環境における情報保護制御機構の開発
キーワード: エッジコンピューティング, 情報保護, 制御機構, コンテナ, 5G, 2020.04～2021.01.
- ・ 地域公共交通への支援活動
キーワード: 地域公共交通, 2020.04～2021.01.

- ・ 安全に利用可能なパーソナルデータ流通基盤の研究
キーワード：パーソナルデータ, 流通基盤, 2020.04～2021.01.
- ・ マルウェア検知システムへのポイズニング攻撃とその対策
キーワード：マルウェア, ポイズニング, 機械学習, 2019.04～2021.01.
- ・ 仮想マシンマイグレーションにおける情報保護制御機構の開発
キーワード：クラウドコンピューティング, マイグレーション, 情報保護, 2017.04～2021.01.

研究プロジェクト

- ・ 5G 時代のエッジコンピューティング環境における情報保護制御機構の開発
2020.04～2023.03, 代表者：内林 俊洋, 九州大学

研究業績

● 学会発表

1. Toshihiro Uchibayashi, Bernady Apduhan, Takuo Suganuma, Masahiro Hiji, A Cloud VM Migration Control Mechanism using Blockchain, Proc. of the 10th International Workshop on Future Computing System Technologies and Applications (FiSTA2020), 2020.07.
2. Shintaro Narisada, Shoichiro Sasaki, Seira Hidano, Toshihiro Uchibayashi, Takuo Suganuma, Masahiro Hiji, Shinsaku Kiyomoto, Stronger Targeted Poisoning Attacks Against Malware Detection, Cryptology and Network Security (CANS 2020), 2020.12.
3. 稲永 健太郎, 高木 秀也, 末吉 智奈佐, 内林 俊洋, コロナ禍における安全安心な地域公共交通利用に向けた ICT 活用の検討, 日本経営システム学会九州・沖縄支部令和 2 年度第 1 回研究会, 2021. 03.

研究資金

● 科学研究費補助金

1. 2020 年度～2022 年度, 基盤研究(C), 代表, 5G 時代のエッジコンピューティング環境における情報保護制御機構の開発.

1.3.5 櫻井 大督

研究内容

視覚的データ解析, トポロジカルデータ解析

研究プロジェクト

- ・ 仮説のオントロジーに基づく可視化
2020.04～2023.03, 代表者: 櫻井大督, 九州大学

研究業績

● 原著論文

1. Daisuke Sakurai, Kenji Ono, Hamish Carr, Jorji Nonaka, and Tomohiro Kawanabe, Flexible Fiber Surface : A Reeb-Free Approach, Topological Methods in Data Analysis and Visualization V, 2020.10.
2. Martin Falk, Guillaume Favelier, Charles Gueunet, Pierre Guillou, Adhitya Kamakshidasan, Pavol Klacansky, Joshua Levine, Jonas Lukasczyk, Daisuke Sakurai, Maxime Soler, Julien Tierny, Will Usher, Jules Vidal and Michal Wozniak, An Overview of the Topology ToolKit, Topological Methods in Visualization – Applications and Software, Springer, in print.
3. Daisuke Sakurai, Takahiro Yamamoto, Visually Evaluating the Topological Equivalence of Bounded Bivariate Fields, Topological Methods in Visualization – Applications and Software, Springer, in print.

研究資金

● 科学研究費補助金

1. 2020年度～2022年度, 若手研究, 代表, 仮説のオントロジーに基づく可視化

● 競争的資金

1. 2020年度～2020年度, 令和2年度 IMI 共同利用研究・研究集会 II, 代表, Fiber Topology Meets Applications

教育活動

- 教育活動概要

汎オミクス計測科学センターでの若手研究者育成，若手技術育成など