

修士論文

2015 年度

自己組織化マップを用いた
プライバシー考慮電力データ共有手法の提案

岡田健吾

(学籍番号：81419193)

指導教員 教授 西 宏章

2016 年 3 月

慶應義塾大学大学院理工学研究科
開放環境科学学専攻

論文要旨

通信機能を持った電力メーターであるスマートメータ導入に伴い、家庭のエネルギーを管理するシステムであるHEMS(*Home Energy Management System*)で得られたデータをサービス提供のために用いるなど、データの二次利用(電力需要のピークカットを目的としたデマンドレスポンスなど)が考えられている。さらに日本では 2016 年度より電力小売り自由化が始まる。これにより各家庭は電力利用パターンに応じて最も適している電力会社から電力を買うことが可能となる。そのため同じ地域の家庭でも異なる電力会社と契約する場合が予測され、電力利用データの点在化が進む。点在化したデータを利用する場合、企業間でのデータシェアする必要があるが、電力利用データはプライバシー情報と考えられており、匿名化処理が必要である。

プライバシーを保護しつつデータを共有する手法は、匿名化を用いるアプローチが考えられるが、それぞれの企業が独自に匿名化を行うと、データ数が少ない場合にデータの損失が大きくなる。本研究では、教師なし学習の SOM(Self-Organizing Map)を用いることで点在化するデータの傾向を考慮して代表パターンを作成し、それぞれの実データと代表パターンの類似度をそれぞれ計算し、最も類似度が高いものを実データの代わりに用いてこのデータを共有する手法を提案した。データを共有することにより、匿名化の際の対象データ数を増加させ、少ない情報損失度で匿名化処理を行うことを可能とした。

本研究では、データ利用アプリケーションとしてデマンドレスポンスを想定しており、実データと代表パターンの類似度の計算の際に需用電力のピーク時刻の類似度に対する重みづけを行った。需用電力のピーク時刻の予測には、時系列解析手法である季節自己回帰和分平均モデル(SARIMA)を用いて、翌日需用電力の予測を行うことでピーク時刻を決定した。

提案手法でデータを共有した後に匿名化を行った場合と、それぞれの企業が匿名化を行った場合のデータの比較を行った。評価の結果、匿名化によるデータ保護レベルの指標の一つである k 匿名性(k が大きい程強い匿名性)の k 値が大きく、そしてデータを共有する企業数が多い程、提案手法を用いた場合の方が少ない情報損失で匿名化処理を行うことができることを示した。特に、 $k=20$ 、企業数が 15 の場合、提案手法を用いることで情報損失度を 20%削減することを達成した。

Thesis Abstract

A smart meter is a novel electric power meter that can be installed in every household. It has a communication function and transmits data obtained from the HEMS (Home Energy Management System) to energy providers. These data are used not only for providing service but also for the peak cut of power demand “Demand Response”. Furthermore, in Japan, electricity deregulation of households is expected to commence in 2016. One of the advantages of electricity deregulation is that residents can select a suitable plan according to their consumption patterns. Once electricity deregulation is introduced, many energy providers will start providing their services, and each household will select a suitable energy provider. Therefore, different households in the same area may agree contracts with different energy providers, so the data are stored in database of different providers. When the company used the all household data, it is needed to collect the power consumption data which other companies. Therefore, the method for privacy preserving data sharing is needed to share the power consumption data which other companies have.

The data sharing method are proposed anonymization methods. However, anonymization methods caused large loss of data when each companies anonymizes own data, because the number of data is small. In this thesis, an unsupervised learning of a SOM (Self-Organizing Map) is used to create a representative pattern considering tendency of the data stored in all providers. The similarity of each of the real data and representative pattern are calculated respectively. The data which has the highest degree of similarity is used instead of the actual data. By sharing the data, to increase the number of target data make it possible to conduct an anonymous process with less information loss degree.

In this thesis, Demand Response is considered as a data use application. The weighting calculation focused on the time of peak demand in the similarity of each of the real data and representative pattern was carried out. Seasonal Autoregressive Integrated Moving Average (SARIMA), which is time series data analysis, is used to predict the time of peak demand power by prediction of the next day demand power.

The anonymization after sharing data using proposed method were compared with when each company anonymizes their own data respectively. The result is that larger k value, which k anonymity is one of the indicators of the level of data protection, and the larger the number of companies that share data lead to small data loss. Information loss was reduced 20% by using the proposed method in case value of k is 20 and the number of provider is 15.

目次

第1章 序論	1
1.1 スマートグリッド	1
1.1.1 スマートメータ	2
1.1.2 デマンドレスポンス	2
1.2 電力小売り自由化	4
1.2.1 概要	4
1.2.2 特定規模電気事業者	5
1.3 個人情報保護とプライバシー	5
1.3.1 米国:消費者プライバシー権利章典(Consumer Privacy Bill of Rights)	6
1.3.2 スマートメータとプライバシー	7
1.4 本研究の概要	8
第2章 関連研究	10
2.1 次世代エネルギー・社会システム実証実験	10
2.1.1 横浜市(YSCP:横浜 スマートシティ プロジェクト)	10
2.1.2 北九州市	12
2.2 小杉駅周辺地区スマートコミュニティ事業	14
2.3 <i>Privacy Preserving Data Mining (PPDM)</i>	16
2.4 プライバシー基準	18
2.4.1 用語の説明	18
2.4.2 k -匿名性	19
2.4.3 l -多様性	19
2.5 k -member clustering	19
2.6 準同型暗号化	21
2.7 Self-Organizing map (SOM:自己組織化マップ)	22
2.8 時系列解析手法	23
第3章 前提条件と提案電力データ共有手法	26
3.1 概要	26
3.2 提案するSOMを用いた電力データ共有手法	27
3.2.1 電力データ	27
3.2.2 SOMを用いたデータ共有アルゴリズム	27
3.2.3 デマンドレスポンスを目的とした重み付け	31
3.3 翌日需要電力のピーク時刻推定	32
第4章 データ取得環境の構築	35
4.1 川崎市スマートシティプロジェクトにおけるシステム	35

4.2	IEEE1888.....	35
4.2.1	概要	35
4.2.2	データの取得	37
4.2.3	データのフォーマット.....	38
第 5 章	評価	40
5.1	動作環境, 対象データ	40
5.2	k-匿名化による情報損失度の比較.....	40
5.2.1	武蔵小杉データセットを用いた評価	41
5.2.2	Irish データセットを用いた評価	41
5.3	デマンドレスポンスを想定した評価	44
第 6 章	結論	47
参考文献		

目 次

図 1-1 スマートグリッド概念図 [1]	1
図 1-2 東京電力のスマートメータ(左)と関西電力のスマートメータ(右)	3
図 1-3 ある一般家庭の電力使用の様子 [7]	8
図 2-1 実証事業における構築されたシステムとイメージ図 [10]	11
図 2-2 統合BEMSによる大規模ビル間連携実証の実施内容 [11]	11
図 2-3 統合BEMSによる大規模ビル間連携実証のピークカットイメージ [11]	12
図 2-4 東田地区のエネルギー需給をモニターする「地域節電所」 [12]	13
図 2-5 「Dr. Driveセルフ八幡東田店」と急速充電システム [13]	14
図 2-6 お出かけでお得, クールシェアイメージ [14]	15
図 2-7 お出かけでお得によるピークシフト [15]	15
図 2-8 医療データと投票者名簿	16
図 2-9 ランダム化アプローチ	17
図 2-10 暗号学的アプローチ	17
図 2-11 匿名化アプローチ	18
図 2-12 一般化階層	20
図 2-13 k - member clustering アルゴリズム [24]	21
図 2-14 自己組織化マップ	23
図 3-1 想定する前提条件	26
図 3-2 待機電力と活動電力の分離	27
図 3-3 学習ステップ	28
図 3-4 カウント情報は秘匿にすべき情報で, 企業が結託する可能性がある場合のカウント 情報集計	30
図 3-5 カウント情報は秘`匿にすべき情報で, 企業が結託しない場合のカウント情報集計	30
図 3-6 ピーク時間が 18 時の場合の重み付け分布	31
図 3-7 自己相関 (ACF) と偏自己相関 (PACF)	33
図 3-8 翌日需要電力のピーク時刻推定の流れ	33
図 3-9 SARIMA における, モデル化と予測イメージ	34
図 3-10 SARIMA を用いた一日先需要予測結果と実データ	34
図 4-1 川崎市スマートシティプロジェクトにおけるシステム全体図	35
図 4-2 IEEE1888 で構成されるシステム [21]	36
図 4-3 Fetch 手順 (通信シーケンス)	37
図 4-4 body クラスの表現	38
図 4-5 全世帯の情報取得時のポイントツリー	38
図 5-1 電力データの k -匿名化	41
図 5-2 k と企業数の関係 MapSize 5×5 (武蔵小杉データセット)	42
図 5-3 k と企業数の関係 MapSize 10×10 (武蔵小杉データセット)	42

図 5-4 k と企業数の関係 MapSize 10×10 (Irish データセット)	43
図 5-5 k と企業数の関係 MapSize 20×20 (Irish データセット)	43
図 5-6 データサイズを変化させた場合の MAE による評価.....	44
図 5-7 武蔵小杉データセットの MAE	45
図 5-8 Irish データセットの MAE	45

表 目 次

表 2-1 医療記録例	20
表 2-2 匿名化後の医療記録例 ($k = 2$)	20
表 5-1 MapSize 5×5 における <i>Ratenp</i> (武蔵小杉データセット)	42
表 5-2 Map Size 10×10 における <i>Ratenp</i> (武蔵小杉データセット)	42
表 5-3 MapSize 10×10 における <i>Ratenp</i> (Irish データセット)	43
表 5-4 MapSize 20×20 における <i>Ratenp</i> (Irish データセット)	43
表 5-5 データサイズを変化させた場合の MAE による評価.....	44

第1章 序論

1.1 スマートグリッド

化石燃料枯渇問題, 地球温暖化問題に伴い, 新エネルギーの開発や省エネ対策としてのネットワーク統合制御システムの開発が積極的に行われてきた. 近年その中でも, 特に次世代電力網である, スマートグリッドが大きく注目を集めるようになった. スマートグリッドとは, 電源・電力流通設備・電力需要家設備にICT(*Information and Communication Technology*)技術や最先端技術を活用し, より環境性・品質性・効率性を高める電力供給システムである [1].

(図 1-1) スマートグリッドは自然エネルギー及び分散電源利用の効率化に寄与し, 電気供給の安定性の向上にもつながるシステムである. 自然エネルギーを利用した発電設備や分散電源などは今後増加すると予想され, これらの小型発電設備を大電力網と接続させることが大きな命題である.

2005 年にEU委員会が次世代の技術開発構想としてスマートグリッドという言葉を発表した. その後, 米国のオバマ大統領が就任して直後の 2009 年にグリーン・ニューディール政策を展開し, スマートグリッドの技術開発に 45 億ドルの予算を投じることを決定した. このプロジェクトには電力会社をはじめ, *Google*, *IBM*などが参加しており, 全米の主力企業を巻き込んだプロジェクトとなった.

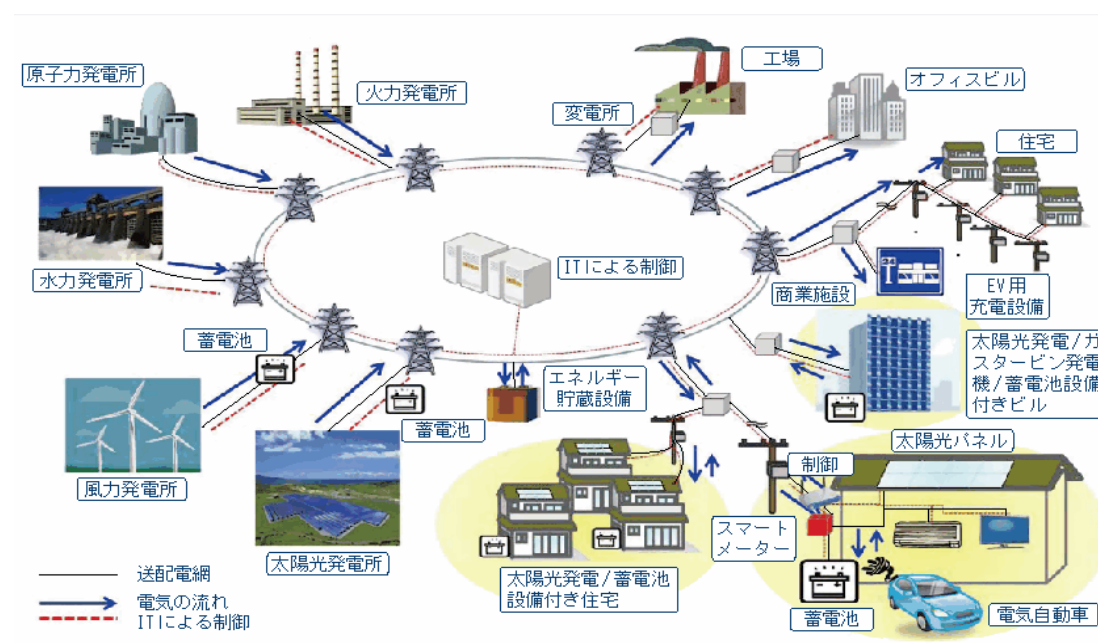


図 1-1 スマートグリッド概念図 [1]

第 1 章 序論

1.1.1 スマートメータ

スマートグリッドでは既存の電力計の代わりに、スマートメータ(図 1-2)という新しい電力計を設置する。スマートメータは通信機能を備えており、ネットワークを通じて消費電力量をリアルタイムで計測することが可能である。既存の電力計では検針員による目視によって月に一度など一定周期に検針が必要であり、リアルタイムでの検針は不可能であった。リアルタイムでの検針が可能であると、料金設定を柔軟に行うことができる。たとえば電力使用量の多い時間帯、あるいは多くなる傾向の時間帯の料金を高くすることで需要者側は節約をするようになり、全体としてはピークカット、個人としては節約につながる。

日本ではスマートメータの普及はあまり進んでいない。2013 年 9 月 11 日に開催された「スマートメータ制度検討会」での報告 [2]によると、500kW 以上の特別高圧、高圧大口の利用者に対しては 8 割で設置が完了し自動検針を実施済みである。一方使用容量の少ない高圧小口や低圧利用者には 2%しか設置されておらず、既存の機械式メーターのままとなっている。電力会社では、関西電力が平成 20 年度より本格的導入に向けた実証実験を開始し、九州電力が積極的に導入を始めている。

次に海外のスマートメータ導入状況について述べる。

アメリカ連邦政府は 2003 年ごろからスマートグリッドに関するレポートなどを発表している。それに伴い、スマートグリッド実現に向けた法律や開発支援のための法律を施行しておりスマートメータの導入を後押ししている。特にカリフォルニア州、ペルシルバニア州、テキサス州がスマートメータの先行導入している [3]。アメリカでは産業・業務需要家を中心に需給調整契約や緊急時デマンドレスポンス(後述)といった、需要家側付加抑制の取り組みに何かしらのインセンティブを与えるデマンドレスポンスの効果が大きいという報告もある [4]。

欧州では2002年のイタリアでの導入に始まり、スペイン、イギリス、フランスなど導入の義務化がなされている国も存在している。導入している国は多数あるが、料金メニューの取り組みも多岐に渡っている。

アジア・オセアニアでは中国・韓国で導入が急速に進んでいる。これに伴って国内の複数のメーカーが価格競争を行っている。中国では 2015 年までに 2.3 億個のスマートメータを導入している。オーストラリアではビクトリア州の取り組みが先行しているが、2011 年末にIHD (*In Home Display*) に対する補助金や柔軟な料金形態の導入延長など制度に対する変更点を発表している。

1.1.2 デマンドレスポンス

デマンドレスポンス(DR: *Demand Response*)とは、電力需要のピークに対して節電に関する、インセンティブや電力価格の変更を、電力会社あるいはアグリゲータから電力利用者に通知し、



図 1-2 東京電力のスマートメータ(左)と関西電力のスマートメータ (右)

需要家が利用を控える, あるいはピーク時以外に利用をずらしたりすることによって, ピーク電力を削減する方法である. 米国では実証実験が行われ 10~20%の需要削減効果が確認された. [5]. デマンドレスポンスは大きく電力料金ベースとインセンティブベースのデマンドレスポンスに分かれる.

1.1.2.1 電気料金ベースのデマンドレスポンス

電気料金ベースのデマンドレスポンスとは, 時間帯別料金を設け, 高需要時には料金を高くして需要を抑制し, 割安な料金に設定された時間帯にピークシフトするように需要家の行動を促進する仕組みである. 具体的なメニューを以下に示す.

- *Time of Use Pricing*

時間帯別の平均卸費用・託送費用を反映し 1 時間から 24 時間の間隔で変動する電気料金.

- *Critical Peak Pricing*

平常時は時間帯別特別料金であるが, 特定条件(緊急時, 卸電力価格高騰時)により高い単価が適用される電気料金.

- *Real Time Pricing*

時間帯別に形成される卸電力市場価格を反映させ, 毎時間あるいは頻繁に変動する電気料金.

- *System Peak Response Transmission Tariff*

インターバルメータをつけ, 料金を含む特定条件化で高需要期に電力需要を削減して託送料金を削減する方法.

1.1.2.2 インセンティブベースのデマンドレスポンス

料金ベースのデマンドレスポンスは電力料金を変えることで需要家側の電力利用の削減を目的としていた. インセンティブベースのデマンドレスポンスは, 電気事業者などの供給者側と, 需

第 1 章 序論

要家側と契約を結び、卸価格が高騰または電力需給が逼迫している際に負荷抑制、遮断を要請または実施する仕組みのことで削減を狙う仕組みである。以下にインセンティブベースのメニューを示す。

- *Direct Load Control*

直前の通知により、顧客のエアコンなどの電気機器を遠隔で遮断もしくは循環運転するように自動制御する手法。

- *Interruptible load*

システムの不測の事態時に、需要家に負荷軽減に同意させる代わりに、料金割引などの提供を契約したうえで、電気の使用量が削減または遮断されるメニュー。

- *load as Capacity Resource*

システムに不測の事態が起こった際に、需要家にあらかじめ規定された負荷削減量を約束させるメニュー。

- *Spinning Reserves*

緊急時の最初の数分間、供給と需要のインバランスの解決策を需要家に同期させいつでもできるようにするプログラム。

- *Non – Spinning Reserves*

すぐには利用できないが、10 分遅れで供給と需要のインバランスの解決策を需要家に提供できるようにするプログラム。

- *Emergency Demand Response*

負荷削減策を課されている間に得られた負荷削減量に応じて、顧客にインセンティブを与えるデマンドレスポンスシステム。

- *Regulation Service*

システム運用者からリアルタイム情報に応じて、需要家が負荷を増減させるプログラム。

- *Demand Bidding and Buyback*

需要家に小売り、卸売市場において、ある価格で負荷軽減を申し出ることや、特定の単価でどれくらいの負荷を進んで削減するかを特定することを許可するプログラム。

1.2 電力小売り自由化

1.2.1 概要

日本では 2016 年度より電力小売りの自由化が始まる。小売りとは契約電力が50kW 以下の低電力帯を扱う部門のことを表す。それ以上の大口の契約電力に関しては 2000 年以降の電力時事業制度改革によって順次自由化が行われ、既に完了している。

第1章 序論

電力の小売り自由化が行われるまでは各地域によって電力を購入できる電力会社は決まっており(一般電気事業者)、電力供給の安定性が保たれていた。しかし独占的であったため、価格競争が行われず、適切なサービス、価格になっているとは言えない。自由化によって電力市場が開放され、消費者は電力を購入する電力会社を選ぶことができるようになり、多くの事業者が電力市場に参入することで、事業者間の競争による価格の低下が期待できることが自由化のメリットとしてあげられる。2016年の自由化される電力市場規模は8兆円程度と想定されており、市場規模は大きい。

1.2.2 特定規模電気事業者

自由化によって参入する事業者のことを、特定規模電気事業者(PPS、新電力)という。経済産業省資源・エネルギー庁によると小売り電気事業登録は2015年12月1日現在、89件の登録が既に完了している。新電力の中には電気と自社の電力以外のサービスをセットとして売り出す事業者や、他の企業と共同したサービスも考えられている。東京電力は、電力料金と携帯電話通信料のセット割引を行うことが決まっており、ソフトバンクが契約窓口となり、店舗網を活用することにより、全国的なサービスを目指している。

博報堂エネルギーマーケティング推進室の調べによると、電力自由化時に電力会社を「変えてみたい」と答えた人は64%存在しており、重要視する項目は75%が料金の安さが最も高い値を示している。他料金メニューの手続きのしやすさ、企業イメージ、などが重要視されている。仮に「変えてみたい」と答えた家庭がすべて乗り換えを行ったとすると、5兆円ほどの金額分の電力料金が動くことになる。

新電力の新規参入の壁として考えられているのが、コストの問題である。電力事業を開始するためにベース電力を確保するための発電所の建設や電力の託送料金、そしてインバランス料金である。インバランス料金とは、30分単位での電力の供給量と需要量の誤差が3%以内に収めなければならないという同時同量性を満たせない場合に、補給的に供給する電力料金で一般電気事業者に対して支払う料金のことである。インバランス料金は一定でなく、季節や時間帯によって変動する。同時同量性は規模が小さい会社は、規模が大きい会社と比べ需要予測が難しく、またインバランス料金の支払いによって経営が成り立たなくなるという懸念がある。そのインバランス料金の発生を防ぐために考えられたのがバランシンググループである。

バランシンググループは同時同量性をそれぞれの企業で満たすのが困難であることが予測され、複数の企業で過不足を満たすことが出来ればインバランス料金が請求されないとする制度である。例えば他の企業が供給の不足であるときに、他の企業に余剰電力がある場合は合計として3%を越さなければインバランス料金が発生しない。このバランシンググループはグループサイズが大きい程スケールメリットを生かすことが出来るため有効である。バランシンググループの欠点としては、代表契約者となる企業に対して、事務作業や責任が集中すること、また各新電力間での同時同量性を達成するために情報の共有が行われるが、それぞれの企業の情報を公開することになること等が考えられている。

1.3 個人情報保護とプライバシー

多種多様なものがデータ化され、蓄積されるようになった。蓄積されたデータを分析すること

第1章 序論

により、新しいビジネス創出などに繋げるなど様々に活用されている。それに伴い、個人の情報の安全な利用を目指した個人情報保護法が施行されているが、IT技術やビジネスの発展が急激に進むにしたがってさまざまな問題が生じている。個人情報保護法において保護対象とされる個人情報の定義がIT技術に対応できない場合（個人の名前は特定できないが、データマイニングにより個人が特定される等）がある事や、情報主体が自分の情報の取り扱いについてコントロールできる権利が政策等で言及されているものの、IT技術やサービスを実現するための具体的な実態が不明瞭であり、技術発展が進むことに対し、法制度が整う速度が追いつかないという問題もある。それに伴いパーソナルデータに関する利活用ルールの明確化等を目的とした、パーソナルデータに関する検討会、そして技術検討ワーキンググループが共に平成25年から行なわれるようになった。

「個人情報」と「プライバシー」は一般的に区別する事なく認識されているが、厳密には異なる概念である。個人情報とは「生存する個人に関する情報であつて、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの（他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む。）」と個人情報保護法第2条1項において定義されている。個人が特定可能であれば個人情報であり本人の意思とは関係なく、本人に関わるすべての情報を指す。一方で、プライバシーは法令上で定義されておらず曖昧である。一般的にプライバシーな情報とは、個人の私生活等、一般に公開されてほしくない情報の事を示す場合が多い。ライフログなど一時的なデータは個人識別に繋がる事はないが、長期的に蓄積され時にプライバシーの問題を引き起こす可能性はある。

1.3.1 米国:消費者プライバシー権利章典(Consumer Privacy Bill of Rights)

アメリカにおけるプライバシー法は包括的な法でなく、分野ごとに個別の法律（電子通信プライバシー法、金融サービス近代化法等）によって定められているが、分野に当てはまらない場合の基本原則を定めたものが消費者プライバシー権利章典である。7原則を以下に示す。

- 個人によるコントロール

消費者は、企業がどのような個人情報を収集し、それらをどのように利用するかについて、コントロールを行う権利を有する。

- 透明性

消費者は、プライバシーやセキュリティ慣行に関して、容易に理解でき、かつアクセスできる情報を有する権利を有する。

- コンテキストの尊重

消費者は、消費者が情報を提供したコンテキストに合致する方法で、企業が個人情報を収集、利用、開示することを期待する権利を有する。

- セキュリティ

消費者は、個人情報を安全かつ責任のある取り扱いをしてもらう権利を有する。

- アクセス及び正確性

第1章 序論

消費者は、当該データのセンシティブティーや情報が不正確であった場合の悪影響のリスクに照らして適切な方法を通じ、利用可能なフォーマットで、個人情報にアクセスし修正する権利を有する。

- フォーカスを絞った収集

消費者は、企業が収集・保持する個人情報について合理的な制限を設ける権利を有する。

- アカウンタビリティ

消費者は、個人情報、消費者プライバシー権利章典を順守するための適切な措置を講じた企業により、適切に取り扱われる権利を有する。

1.3.2 スマートメータとプライバシー

スマートメータは上述したようにスマートグリッドを成す上で、重要な要素の一つである。一方、プライバシーの観点で問題が出てくる可能性が懸念されている。スマートメータは電力の使用状況をリアルタイムで送ることが可能であるが、このデータから家庭環境が分析される可能性がある。*Smart Grids and Cleanpower Conference* では「我が社は、スマートメータを利用して詳細に電力消費データを収集、記録することが出来る。この性能をもってすれば、家庭内でそれぞれ何人住んでいて、何をしているのか、一階にいるのか二階にいるのか、犬を飼っているのか、何時に起きるのか、シャワーを何時に浴びるのか、等、膨大な量のプライバシー情報を吸い上げる事が可能である」と述べている [6]。

Smart Grid Interoperability Panel – Cyber Security Working Group と呼ばれる NIST 組織の一つは、ある一般家庭の電力使用の様子 (図 1-3) に関して分析を行った。電力使用の様子から、以下のような事を推定することが出来てしまうと報告している [7]。

- 冷蔵庫の稼働状況 (どの程度の大きさなのか=住居者人数)
- お湯を沸かす、トーストを焼いている時間 (朝食時間)
- 洗濯機を回している時間 (確実に家事を行って、家に大人がいる時間帯)
- オーブンを使っている時間帯 (夕食時間)
- 夜にお湯を沸かしている時間 (家族がくつろいでいる時間)

電力使用状況を分析することで、このように家庭内の状況を想定することが可能である。家庭内の状況というのは通常情報化されないもので、一般的な個人データよりもプライバシー保護の必要性が高いとも言える。仮にこのデータが公開されるとすると生活パターンが分かるので、例えば保険会社の保険料のシステムに影響を与える、さらには、広告会社は生活パターンに合わせた広告の提示をするなど、新たなサービス展開も考えられる。悪用することを考えるのであれば電力をあまり消費していない時間帯、つまり留守である可能性が高い時間帯を狙って空き巣に合う可能性がある。このように電力使用データは、電力会社が各家庭の電力料金を決めると

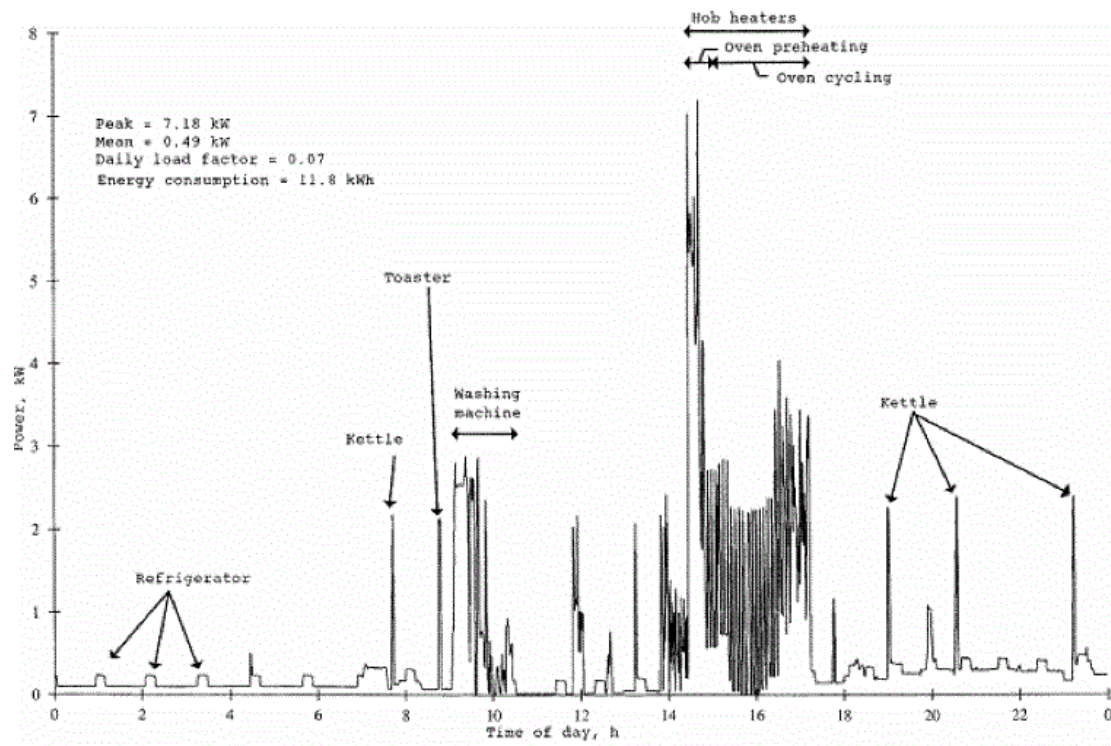


図 1-3 ある一般家庭の電力使用の様子 [7]

いう目的だけではなく、二次的利用も考えられるので、データの扱いには慎重になる必要があると考えられる。

情報の二次利用を考える際、プライバシーを保護する必要がある。そこで提案されたのがプライバシーを保護しつつデータマイニングを可能とするプライバシー保護データマイニングや、データに対して加工を施しプライバシーを保護した状態で公開するデータパブリッシングである。これらの技術は個別の情報を明らかにすることなく、データを公開する、もしくはデータから相関関係や統計情報などの有用な情報を抽出する方法である。プライバシーを保護するためにデータを削除、抽象化しなければならないが、情報の有用性を保つためになるべくデータを保持しなければならない。このようにトレードオフの関係であるプライバシー保護と情報の価値のバランスを適切に設定し、価値のある情報を完全に活用するための技術が研究されている [8]。

1.4 本研究の概要

スマートメータ導入に伴い、家庭のエネルギーを管理するシステムである HEMS(Home Energy Management System)で得られたデータをサービス提供のために用いるなど、データの二次利用が想定されている。サービス展開だけではなく、各家庭の電力の利用状況から、電力需要のピークカットを目的としたデマンドレスポンスを適切に行うことも可能となる。

電力使用量にはデータとしての価値があるが、電力使用量を分析することにより家庭内の様子が推定されてしまうといった問題があり、プライバシーの観点から、データの第三者提供には

第 1 章 序論

匿名化処理が必要である。また、日本では 2016 年度より電力小売り自由化が始まるが、同じ地域の家庭でも異なる電力提供企業と契約することが考えられる。スマートメータのデータは契約している企業へ送信されるため、同じ地域の家庭のスマートメータデータは分散的に保存されるようになる。しかし、デマンドレスポンスなど地域として電力利用状況を把握するためにデータを共有する手法が必要である。

以上から本研究では、プライバシーを保護しつつ分散保存されるスマートメータデータを共有する手法を提案する。提案手法は教師なし学習である自己組織化マップをベースとした手法で、実データを公開せずパターンデータの共有を行うことで電力利用を把握する。データを共有した後に匿名化を行った場合とそれぞれの企業が匿名化を行った場合のデータの比較を行い、またデマンドレスポンスの計画に利用するパターンデータと実データの差の評価を行う。

本論文では第 2 章で関連研究を述べる。第 3 章では提案する共有手法について述べる。第 4 章は評価のためのデータ取得環境について述べる。第 5 章では提案手法の評価と考察を行い、第 6 章で本論文の結論を述べる。

第2章 関連研究

2.1 次世代エネルギー・社会システム実証実験

現在、「次世代エネルギー・社会システム実証実験」[9]は、横浜市、豊田市、けいはんな学研都市(京都)、北九州市の4地域で行われており、実証実験は2010年から2014年までの5か年計画である。経済産業省は2009年11月に省内横断的なプロジェクトチーム「次世代エネルギー・社会システム協議会」を設置した。2010年1月には、次世代のエネルギー流通および社会システムのあり方に対する中間的な取りまとめを発表した。この取りまとめを受け、スマートグリッド/スマートシティの社会実証地域が公募され、選ばれたのが先ほどの4都市である。この実証実験の目的は日本におけるスマートグリッド、スマートシティのあり方を見出すのが目的である。スマートグリッドの関連技術の発展だけではなく、仕組みや、ビジネスモデルの検証も目的に含まれている。他にも、エネルギーの見える化、エネルギー需要に応じて供給者側から需要調節を促すデマンドレスポンス、EV、蓄電システムの設計などを行っている。

2.1.1 横浜市(YSCP:横浜 スマートシティ プロジェクト)

横浜市では、すでに社会インフラが整備されている都市を、快適かつ低炭素な都市に変革することを目的としている。地域のエネルギー需給を統合的に管理する**CEMS** (*City Energy Management System*)を中心とした地域エネルギーマネジメント・システムの開発・運用とPVなどの普及促進などを計画している。実証は日本最大級の複合都市「みなとみらい21地区」、7万5000世帯以上が暮らしベッドタウンとして成長を続ける「港北ニュータウン」、工場や産業団地・公共施設などが集積する臨海部「横浜グリーンバレー」という、地域が異なる3エリアにて展開されている。太陽光発電をはじめとする再生可能エネルギーを大量導入し、2000台の電気自動車EV、4000件の家庭のエネルギーを統合的に管理するシステムである**HEMS** (*Home Energy Management System*)を導入している。ガス・電力の使用状況の見える化のみではなく、電力・熱エネルギーの統合制御も実証している。

具体的な実証内容は、市内にある130世帯に**HEMS**を導入するというものである。これは、横浜市内のマンション管理を受託している大京アステージが提案したものである。また、温湿度照度センサーなどの**HEMS**関連機器の導入も同社によって行われている。各家庭の電力使用データは30分毎に大京アステージのサーバーに集約され、YSCPの**CEMS**と連携している。一時間の電力使用量やデマンドレスポンスに対応して付加されたポイント(インセンティブ)などの情報がパソコン、タブレット等で確認できる。構築されたシステムのイメージを図2-1に示す。

エネルギーの見える化だけでなく、**HEMS**を通じて収集した家庭のエネルギーデータを解析することによって、マンションの居住者向けに適したサービスの提供を目指している。例えば、セキュリティサービスでは、電力使用量を監視し、電力使用量が多い時や逆に使用量が全くない時にはメールなどで警告する。他にもトイレに設置した人感センサーがある一定時間以上反応がない際には、室内のインターフォンで警報音を鳴らす、セキュリティ担当者が駆けつけるといった処置をとる。生活サポートサービスでは、エアコンの電力使用量をもとにエアコンの掃除サー

＜構築されたシステムとの連携イメージ＞

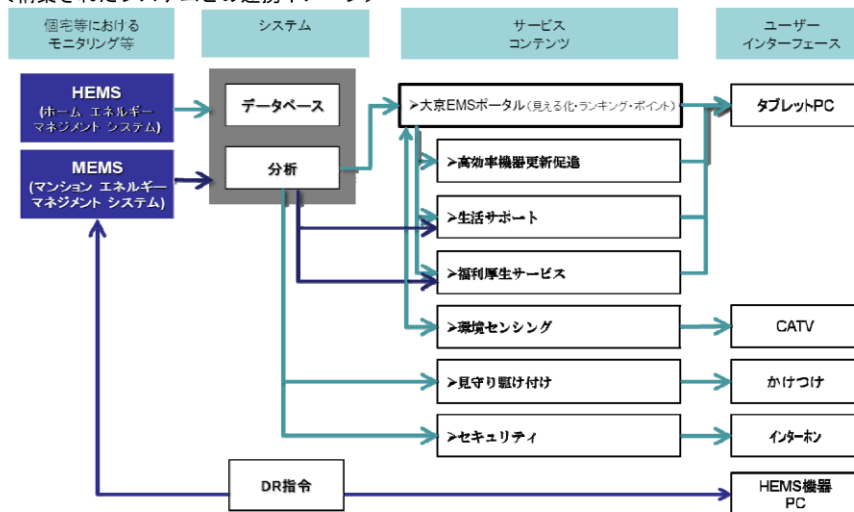


図 2-1 実証事業における構築されたシステムとイメージ図 [10]

■ 実施内容

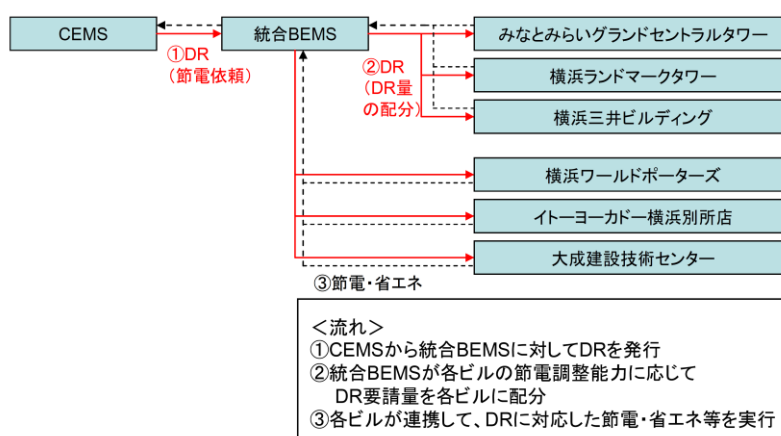


図 2-2 統合BEMSによる大規模ビル間連携実証の実施内容 [11]

ビスを、電力不足時には掃除機に代わるレンタルモップなどをそれぞれリコメンデーションする。福利厚生サービスとしては、外出を促す目的からデマンドレスポンス要請と連動して、外食店などの各種サービス業の割引クーポンをホームページに表示する。このようなHEMSデータを活用して、多様なサービスが今後展開を目指し、大京アステージは「家庭向けサービスの利用を通じた既築集合住宅におけるエネルギーマネジメントの実証」という事業名で実証実験を進めている[10]。

YSCPでは、ビルなどの建物の使用エネルギーを統合的に管理するシステムであるBEMS(Building Energy Management System)も導入している。導入規模は160m²である。CO₂の排出量の見える化機能により、設備区分ごとの電力・ガス使用量を計算し、管理者が機器の制御することでCO₂の削減を目指している。さらに大口需要家であるビルをビル群として管理することで、デマンドレスポンスの対応能力の拡大させる役割を果たすものと予測している。

■ピークカットイメージ

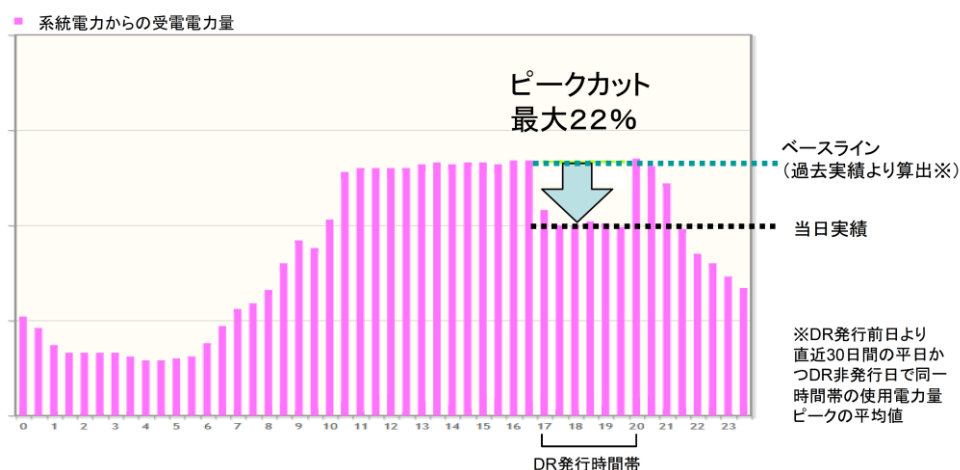


図 2-3 統合BEMSによる大規模ビル間連携実証のピークカットイメージ [11]

2013 年 1 月より大規模ビル 6 か所を対象としデマンドレスポンスの本格運用が始まった。図 2-2 はその連携実証内容を示している。1 月の実証実験のデマンドレスポンスの対象時間帯は 17 時～20 時(平日)であり、その発行条件は、最高気温が前日の夜予報で 8 度以下の日であった。BEMSを活用した業務・商業ビル部門での実証として、電力ピークカットの効果の最大化を目的としたデマンドレスポンスに成功し、最大 22% のピークカットを達成した [11]。図 2-3 はピークカットの様子を示しており、デマンドレスポンス発行時間帯の当日実績値(電力消費量)は過去の実績から算出したベースラインよりも低い値となっており、電力消費量の削減を達成している。

2.1.2 北九州市

北九州市は 1901 年に官営八幡製鉄所が創業して以来、重化学工業地帯として発展してきている都市で近代産業の発祥の地と呼ばれていた。60 年代には大気汚染などの問題に直面したが、公害対策・防止に力を入れて問題を乗り越えてきた。「八幡が変われば日本が変わる、日本が変われば世界が変わる」というスローガンを掲げている。

対象エリアは北九州市八幡東区東田地区の約 120haであり、5 年間で総事業費 163 億円である。実証実験では新エネルギーの導入の強化や地域エネルギーマネジメント、交通システムの整備などにより現状より 20%の削減を目指している。実証実施にあたり以下の 8 つが基本的な考えとなっている。

- 期間エネルギーと新エネルギーの融合による「地域エネルギー共有社会」
- 「地域節電所」を通じた地域エネルギーの「全体最適と部分最適の両立」
- ライフスタイル、ビジネススタイルの変革を促すエネルギーの「見える化社会」
- 市民もエネルギーマネジメントに参加する「エネルギーコミュニティーの構築」
- 上述のシステムに立った都市インフラ、交通など、「都市システム」の整備



図 2-4 東田地区のエネルギー需給をモニターする「地域節電所」 [12]

- 個別技術はもとより社会実装技術を包含した「社会システム技術の開発」や「ビジネスモデル・雇用の創出」
- 国際標準を視野に入れた「世界標準となるモデルの構築・発信」
- 上述事項をパッケージ化して「アジア地域への移転体制の構築」

このような考えのもとに地域エネルギーマネジメントや低炭素社会のあるべき姿を具現化した成果をアジア中心とする全世界に展開しようとしている。

同事業の中核施設といえるのが、「地域節電所」である [12]。これは大規模施設ではなく、図 2-4 のようにオフィスビルのワンフロアの一面に管理システムが置かれているだけである。地域節電所は *HEMS* や *BEMS* などがネットワークで結ばれ、一般家庭 230 戸と企業 70 社の電力需要を監視・予測する。ダイナミックプライシングという電力需給状況に応じて電力料金を日々変化させることで、消費行動に変化を促し、一般家庭の反応の検証が行われている。中でもきめ細やかな負荷平準化を目的とした、「リアルタイムプライシング」という前日の電力需要予測と再生可能エネルギー発電量予測をもとに地域電力の需給状況に応じて当日の電力料金を変動させる仕組み（価格は 2 時間前までに通知）を世界で初めて導入した。

家庭の電力料金だけではなく、2012 年末から電力需要に応じて発令されるダイナミックプライシング情報に合わせて *EV* の充電サービス料金も変動させる実証が始まっている [13] (図 2-5)。ところが、国内で展開されている *EV* 充電設備の利用料金は一般に、無料あるいは利用回数による課金であるため、*EV* の充電タイミングをドライバーに考えてもらえるだけのインセンティブを用意するのが難しい。そこで、実験参加の 9 台の *EV* を対象に、ダイナミックプライシング発令時には充電サービス料金を通常時の 7 倍にまで高めることで行動変革を促すといった実証実験をおこなっている。地域内の電気料金が安い時は、安価な料金によってサービスステーションへの来訪を促し、逆に高い時は、充電サービス料金を高く設定することでサービスステーションでの充電需要を抑える。この実証実験を行うために、JX 日鉱日石エネルギーは「*JX EMS*」と呼ばれるシステムを新規に開発した。*CEMS* と連携し、充電サービス料金を変動させる。*JX EMS* は同時



図 2-5 「Dr. Driveセルフ八幡東田店」と急速充電システム [13]

に、サービスステーションが保有する蓄電池の充放電を制御している。基本的には、地域の電力料金が安いときに蓄電池を充電し、高いときに蓄電池から放電する。その際、急速充電器の使用状況を考慮するのが特徴である。

2.2 小杉駅周辺地区スマートコミュニティ事業

都市型住宅などが集積する小杉駅周辺地区の特徴を踏まえ、市民参加型の省エネ行動などエネルギー利用の最適化などの取り組みを通じ、エネルギーを無駄なく効率よく利用する新しい暮らしの形としてのスマートライフの普及と低炭素社会への貢献、暮らしの利便性の向上などに資する持続可能な取り組みとして、事業を展開している。事業は 2014 年より始まり、対象世帯数は約 150 件で、既存街区での取り組みである。事業主体は川崎市、大和総研ビジネスインベーションで、事業費用は川崎市からの委託費で賄っている

本事業で行われているビジネスを以下に示す

- HEMS データに基づく省エネアドバイス(電気使用レポート)

HEMS データを使つての省エネアドバイスのビジネス検証のために行われている。電力は子ブレーカー単位で取得し、前月比、前年比と比べて家電の利用にどのような変化があったかという今月の傾向を示している。今月の傾向で赤字(電力利用増加)になっている家電の利用に関するアドバイスを示すことで、来月の行動へ繋げることを目的としている。空調や換気扇の単位時間当たりの電力使用量が増えている場合、汚れ・ほこりによって効率が悪くなっていると推定し、掃除すること、あるいはクリーニングサービスの利用を推薦するような仕組みとなっている。

- 地元商店街と連携した実証イベント「お出かけでお得」(地域活性化ビジネス) [14] [15]

【お出かけでお得（イメージ）】



【クールシェアによるエネルギー節約（イメージ）】

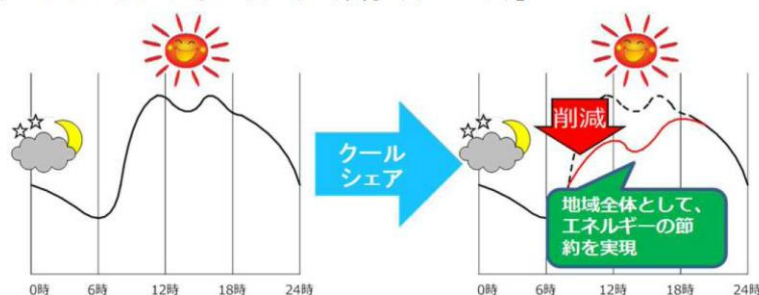


図 2-6 お出かけでお得, クールシェアイメージ [14]

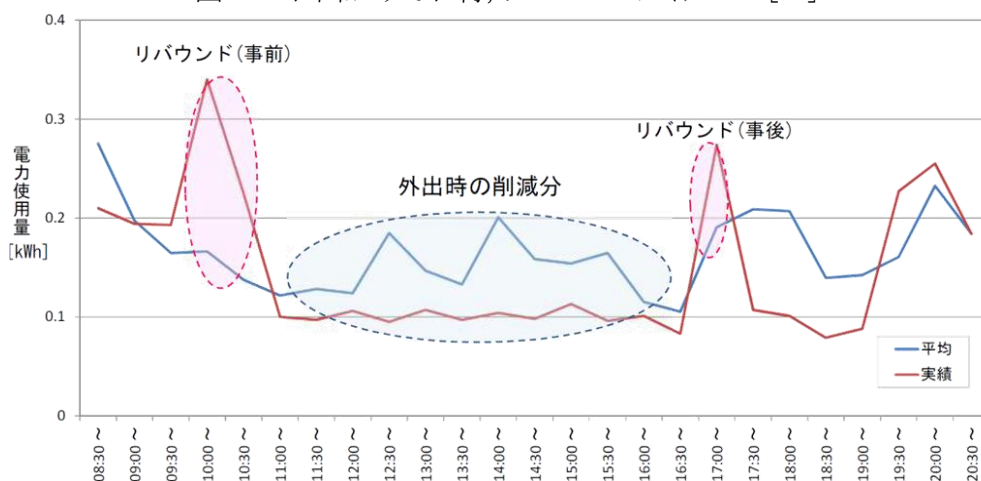


図 2-7 お出かけでお得によるピークシフト [15]

夏における家庭でのエアコンの使用を抑えるため、家庭のエアコンを運転させずに外出し、公共施設など空調が効いた場所に地域の人々が集まることをクールシェアという(図 2-6)。夏の省エネ活動であるクールシェアをトリガーとしてモニターに外出を促すことで地元商店街の回遊につなげ購買規模や省エネ効果を検証する「お出かけでお得」が実施された。図 2-7 はクールシェアに参加した家庭の電力利用を表している。通常の電力利用(平均)と比べ、クールシェアに参加した家庭は日中の電力利用が減っている。一方、外出時間で行っていた可能性がある家事などによる電力利用がリバウンドとなり電力利用が一時的に増えている。しかし、電力需要のピークシフトを行うという観点では有用である。

対象住宅のうちのクールシェアへの参加率は約 10%であった。お出かけによりピークシフトが可能となり、さらに参加者のうちの 1/3 は電力利用量の削減を伴っており、約 12%電力利用を削

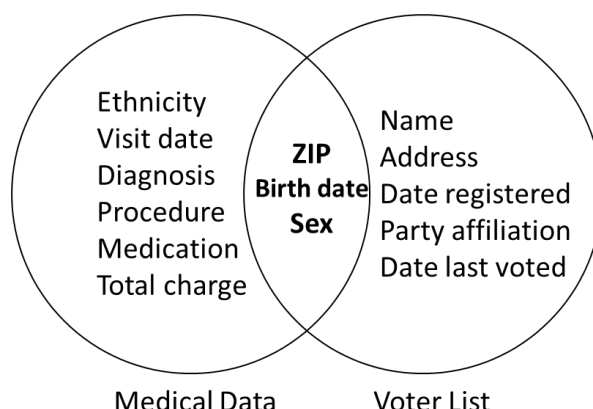


図 2-8 医療データと投票者名簿

減した。この削減量は、外出時間帯での電力利用に加えその前後に消費されている在宅時の電力利用分も差し引き考慮しての削減量である(リバウンドと定義している)。購買状況はお出かけ一回当たり約 1200 円であり、商店街の活性化にも繋がっている。

- 見守りサービス

電気の使用量から高齢者の生活の見守りサービスが実施されている。電気使用量から生活行動を把握し、一定期間電力の使用がない場合にご家族に通知する仕組みである。通常の起床時間を過ぎて、計算された閾値を超える電力利用がない場合に、体調不良などの懸念が通知されるものとした。

高齢者見守りサービスと同様に子供向けの見守りサービスも行われている。平日学校を終えた帰宅時間における子供の帰宅の有無を推定する。帰宅時間が遅かった場合に家族が子供に連絡を取るなどの利用方法が考えられている。

2.3 Privacy Preserving Data Mining (PPDM)

*Privacy Preserving Data Mining*とは、複数の分散したデータベースにおいてそれぞれの機密を保ったまま、つまり機密情報を自身以外に公開せずに、相関関係や統計情報といった有用な情報を得るための技術である [16]。

5 ケタの郵便番号、誕生日、性別の情報を組み合わせると、マサチューセッツ州知事のメディカル情報をユニークに特定することができるという報告がある [17]。マサチューセッツ州では医療データを、氏名を削除した形で公開しており(図 2-8 左円)、一方、選挙の投票者名簿は氏名を公開している(図 2-8 右円)。この二つのデータをマッチングさせると知事と同じ生年月日は 6 人であり、そのうち 3 人の性別が男性であり、郵便番号から 1 人に特定される。公開されているデータを組み合わせることで、プライバシーデータを特定される可能性があるということを意味している。

プライバシーを保護するための研究が盛んに行われている。公開鍵暗号法技術を用いてデータの保護を考えた手法が以前は多かったが、近年では異なるアプローチも多く提案されるようになった。アプローチの例を以下に示す。

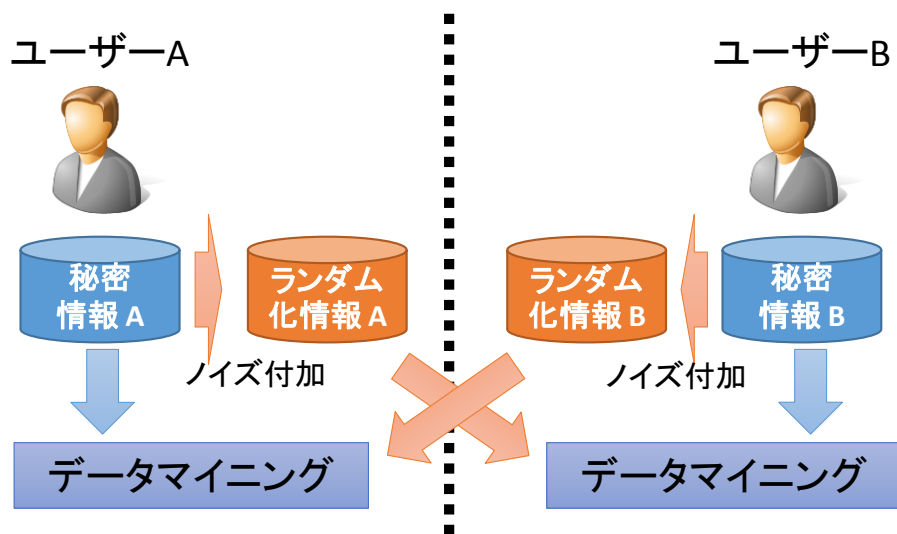


図 2-9 ランダム化アプローチ

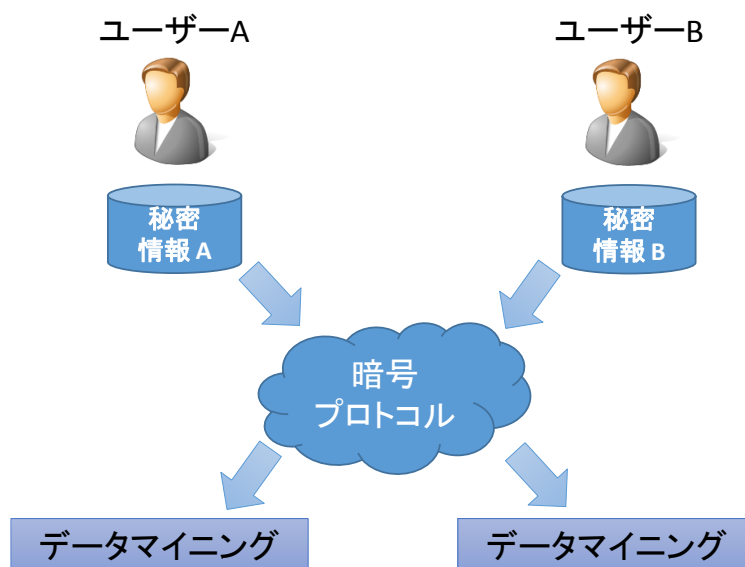


図 2-10 暗号学的アプローチ

- ランダム化アプローチ

秘密情報である元データに対して、任意に設定したノイズを付加したランダム化情報をデータマイニングに利用する手法である(図 2-9)．データマイニングの際にはノイズの確率分布の情報をもとに元データを推定する．元のデータの正確な値を知ることが出来ない状態で、統計情報を分析すると元のデータに類似したものを発行することが可能である．しかし、ランダムノイズの作成方法によって分析結果は変動するので、適切な選択が必要である．さらに、元データが推定されるという点で、安全の保障がされていると言えない．

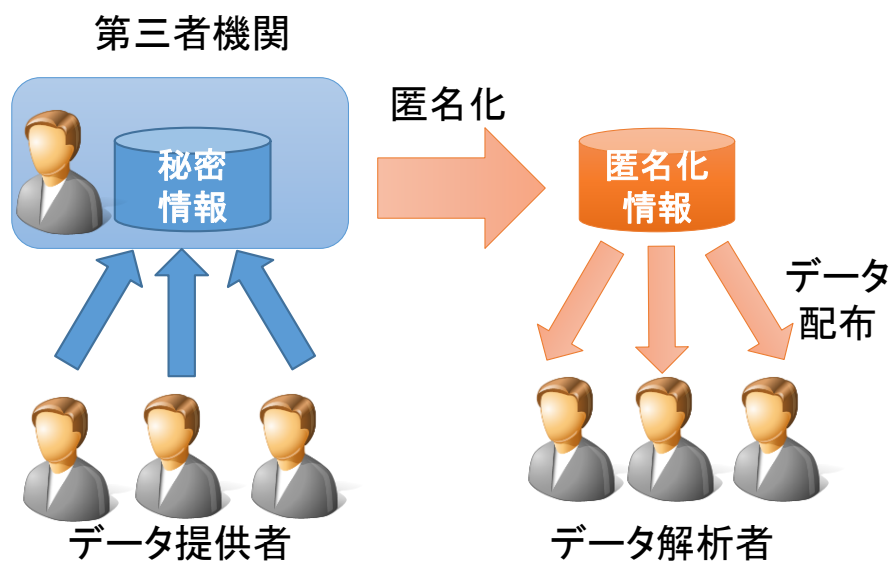


図 2-11 匿名化アプローチ

- 暗号学的アプローチ

データマイニングを複数の計算に分割し、それぞれの計算を安全に行うために暗号化ツールを組み合わせる手法である(図 2-10)。準同型性公開鍵暗号により、暗号化されたデータどうしの加算や乗算を組み合わせた計算が可能である。しかし、この手法は計算量が多いというデメリットがある。

- 匿名化アプローチ

データから一部の情報を削除し、情報量を削減することでプライバシー保護を行う手法である(図 2-11)。匿名化の手法としては「一般化」と「抑圧」がある。一般化とは抽象度によって階層化されているデータであれば上の階層のデータを置き換える操作のことを表す。たとえば、「横浜市」というデータであれば「神奈川県」と置き換える。このようにすることでデータの持つ情報量を削減する。抑圧とは特異性のあるデータを削除する操作のことを表す。特異性のある情報は一般化を行ってもユニークであり、プライバシーが保護できないので削除する。

2.4 プライバシー基準

本論文ではプライバシー保護のレベルを目的に応じて変化させることができるというメリットから、プライバシー保護の手法として匿名化を採用する。ここでは特に匿名化手法におけるプライバシー基準として提案されている、 k -匿名性、 l -多様性に関して述べる。

2.4.1 用語の説明

第 2 章 関連研究

本論文で使用する用語を以下のように定義する.

- データテーブル

本論文で想定するデータテーブルは, 行を「タプル」, 列を「フィールド」と定義する. 各フィールドを「属性」と呼び, 保持するデータ値の意味カテゴリを表す.

- 属性

属性のうち直接プライバシー情報を特定できるものを「識別子」, 識別子ではない属性であっても他の識別子でない属性と組み合わせることによって直接プライバシー情報を特定できる可能性がある. このような属性を「準識別子」と呼ぶ.

- 注目属性

識別子である属性の中でデータマイニングにとって重要な項目であるために一般化を行わない識別子の総称. 一方, 一般化してもよいものを総称して非注目属性と呼ぶ. 一般に非注目属性は準識別子と同一の意味を表す. また, 準識別子が同一であるタプルの集合を q^* -ブロックと呼ぶ.

2.4.2 k -匿名性

データテーブルのプライバシー基準のひとつとして, k -匿名性が提案されている [17]. k -匿名性の定義は, 「あるテーブルにおいて, 同一の準識別子の組み合わせのものが k 個以上存在する状態」である. つまり, q^* -ブロックに k 個以上のデータが存在する状態のことである.

データテーブルの例として表 2-1 を示す. 表 2-1 の郵便番号, 誕生日, 性別を準識別子, 病気を注目属性とする. このテーブルが 2-匿名性 ($k = 2$) を満たすようにデータに抽象化を施し, 匿名化を行うと表 2-2 のようになる. 今回行った匿名化は一般化であり, 数字にはマスキングとして '*' を施し, {'女', '男'} は '人間' というように一段階上の階層に抽象化している. $t1 \sim t3$ の一般化の階層を図 2-12 に記す. 表 2-2 の準識別子の組み合わせが一致しているものがそのデータ自身を含めて 2 つ以上あるので, このテーブルは 2-匿名性を満たしているという.

2.4.3 l -多様性

l -多様性とは [18] において提案されており, 定義は「データテーブルの中のすべての q^* -ブロックにおいて, 注目属性の値が少なくとも l 個の多様性を持っている状態」である. これは背景知識攻撃に対しての問題を解決するためのプライバシー基準であるといえる. 表 2-2 では各 q^* -ブロックにおいて注目属性の種類は最低 2 種類以上あるので, このテーブルは 2-多様性を満たしているといえることができる.

2.5 k -member clustering

データの有用性の失わずに, k -匿名性を満たすための q^* -ブロックの作成を行うことは, データのクラスタリング問題に帰着でき, 現在研究が行われている [20] [21]. k -member clustering という手法が提案されており, この手法は k -匿名性を満たすようなクラスタリングのうちの一つであ

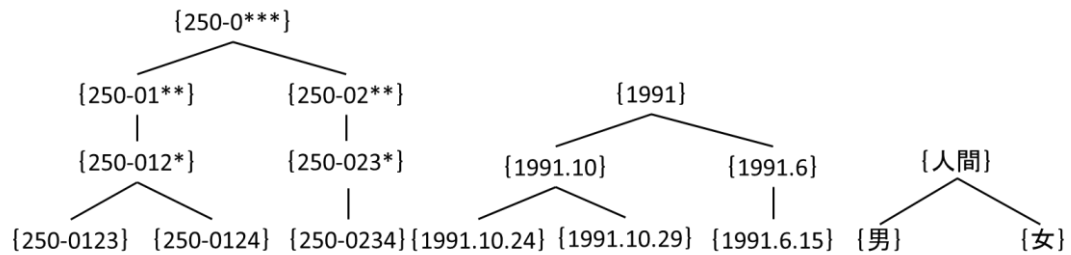


図 2-12 一般化階層

表 2-1 医療記録例

ID	郵便番号	誕生日	性別	病気
t_1	250-0123	1991.10.29	女	癌
t_2	250-0124	1991.6.15	女	風邪
t_3	250-0234	1991.10.24	男	肺炎
t_4	223-1110	2004.8.12	男	風邪
t_5	223-1111	2004.6.17	男	肺炎

表 2-2 匿名化後の医療記録例 ($k = 2$)

ID	郵便番号	誕生日	性別	病気
t_1	250-0***	1991	人間	癌
t_2	250-0***	1991	人間	風邪
t_3	250-0***	1991	人間	風邪
t_4	223-111*	2004	男	肺炎
t_5	223-111*	2004	男	風邪

る. 情報損失が少なくなるように, 各データの距離を定義し, それぞれのクラスタ内の距離の合計値を足し合わせ, その結果が小さくなるようにしている. その合計値が少ないということは情報としての損失量が少ないことを表している.

アルゴリズムは以下のようにになっている(図 2-13).

1. ランダムに初期点を取り, その点から最も遠い点 A を取る.
2. 点 A に近い順に k 個データを集めそれをクラスタ1とする,
3. クラスタ1から最も遠い点を次のクラスタの始点とし, その点に近い点を k 個あつめ, その集合をクラスタ2とする.
4. 手順 3 を繰り返し残りの点が $(k-1)$ 以下になったとき, それらの残りの点はその時点で存在しているクラスタに分配される. 点の分配の方法は, それぞれのクラスタの情報損失の合計値が最も小さくなるようにする.

Function <i>greedy_k_member_clustering</i> (<i>S</i>, <i>k</i>) Input: a set of records <i>S</i> and a threshold value <i>k</i> . Output: a set of clusters each of which contains at least <i>k</i> records. 1. if(<i>S</i> ≤ <i>k</i>) 2. return <i>S</i> ; 3. end if; 4. result = ∅; <i>r</i> = a randomly picked record from <i>S</i> ; 5. while(<i>S</i> ≥ <i>k</i>) 6. <i>r</i> = the furthest record from <i>r</i> ; 7. <i>S</i> = <i>S</i> − { <i>r</i> }; 8. <i>c</i> = { <i>r</i> }; 9. while(<i>c</i> < <i>k</i>) 10. <i>r</i> = <i>find_best_record</i> (<i>S</i> , <i>c</i>); 11. <i>S</i> = <i>S</i> − { <i>r</i> }; 12. <i>c</i> = <i>c</i> ∪ { <i>r</i> }; 13. end while; 14. result = result ∪ { <i>c</i> }; 15. end while; 16. while(<i>S</i> ≠ 0) 17. <i>r</i> = a randomly picked record from <i>S</i> ; 18. <i>S</i> = <i>S</i> − { <i>r</i> }; 19. <i>c</i> = <i>find_best_cluster</i> (result, <i>r</i>); 20. <i>c</i> = <i>c</i> ∪ { <i>r</i> }; 21. end while; 22. return result; End;	Function <i>find_best_record</i> (<i>S</i>, <i>c</i>) Input: a set of records <i>S</i> and a cluster <i>c</i> . Output: a record <i>r</i> ∈ <i>S</i> such that <i>IL</i> (<i>c</i> ∪ { <i>r</i> }) is minimal. 1. <i>n</i> = <i>S</i> ; min = ∞; best = null; 2. for(<i>i</i> = 1... <i>n</i>) 3. <i>r</i> = <i>i</i> -th record in <i>S</i> ; 4. diff = <i>IL</i> (<i>c</i> ∪ { <i>r</i> }) − <i>IL</i> (<i>c</i>); 5. if(diff < min) 6. min = diff; 7. best = <i>r</i> ; 8. end if; 9. end for; 10. return best; End; Function <i>find_best_cluster</i> (<i>C</i>, <i>r</i>) Input: a set of clusters <i>C</i> and a record <i>r</i> . Output: a cluster <i>c</i> ∈ <i>C</i> such that <i>IL</i> (<i>c</i> ∪ { <i>r</i> }) is minimal. 1. <i>n</i> = <i>C</i> ; min = ∞; best = null; 2. for(<i>i</i> = 1... <i>n</i>) 3. <i>c</i> = <i>i</i> -th cluster in <i>C</i> ; 4. diff = <i>IL</i> (<i>c</i> ∪ { <i>r</i> }) − <i>IL</i> (<i>c</i>); 5. if(diff < min) 6. min = diff; 7. best = <i>c</i> ; 8. end if; 9. end for; 10. return best; End;
--	---

図 2-13 *k* − member clustering アルゴリズム [24]

2.6 準同型暗号化

暗号化は様々なものが研究されているが、本節ではセキュアマルチパーティー計算として注目されている準同型暗号について説明する。準同型暗号は平文 m_1, m_2 を暗号化した二つの暗号文 c_1, c_2 ($c_1 = \text{Enc}(m_1), c_2 = \text{Enc}(m_2)$)において、演算子を \circ としたとき、 $c_1 \circ c_2$ の計算結果を復号すると、 $m_1 \circ m_2$ と等しくなる性質を満たす暗号である。特に \circ が $+$ (加法演算子)である場合の暗号を加法準同型暗号という。ここでは加法準同型暗号である Paillier Paillier 暗号 [19]について述べる。具体的なアルゴリズムを以下に示す。

- 鍵生成

2つの大きな素数 p, q をランダムに生成し、 $n = pq$ を計算する。

k をランダムに選び、 $g = (kn + 1) \bmod n^2$ を計算する。

公開鍵を (n, g) 、秘密鍵を (p, q) とする。

- 暗号化

送信するメッセージが m であるとき、以下の式を用いて暗号文 c を生成する。

$$c = \text{Enc}(m) := g^m \cdot r^n \pmod{n^2} \quad (2.1)$$

ここで r は暗号毎に選ばれる n 未満の乱数である。

- 復号

メッセージ m の復号処理は以下の式を用いて行う. $p-1$ と $q-1$ の最小公倍数を λ とした時,

$$m = Dec(c) := \frac{L(c^\lambda \bmod n^2)}{L(g^\lambda \bmod n^2)} \bmod n \quad (2.2)$$

関数 L は式 (2.3) で定義される.

$$L(u) = \frac{u-1}{n} \quad (2.3)$$

- 加法準同型性

平文 m_1, m_2 に対して, 暗号文をそれぞれ c_1, c_2 とする. c を

$$c = c_1 \cdot c_2 \bmod n^2 \quad (2.4)$$

とした時, Parlier 暗号においては式 (2.5) が成り立つ

$$Dec(c) = m_1 + m_2 \bmod n \quad (2.5)$$

2.7 Self-Organizing map (SOM: 自己組織化マップ)

T.Kohonen によって提案された教師なし学習ニューラルネットワークである [20]. SOM は大脳皮質の視覚野をモデル化した学習であり, 図 2-14 のように高次元の入力データ群(ベクトル)をそのデータの関係を保ったまま低次元へと写像することが可能である. 多変量データを高精度かつ効果的にパターンを分類できるものとして SOM は注目を集めている.

入力データを反映したパターン群(ノード)を作成し, トポロジー保存性を有した状態(近いデータほどマップ上で近くに集まる)で多次元ベクトルから二次元へと写像する. 一般的に次元削減, 類似度の計算に用いられる手法である. SOM のアルゴリズムについて述べる.

1. 入力データを反映したパターンデータ(マップ空間中のノード)を生成
マップ空間の初期ノードを準備
2. 入力ベクトル x_i に最も類似度が高いマップ空間中のノードを **BMU**(Best Matching Unit)として選択
3. マップ空間中の **BMU** と周辺ノードの値を入力ベクトルに近づける(BMU に近い程学習の影響を受ける)

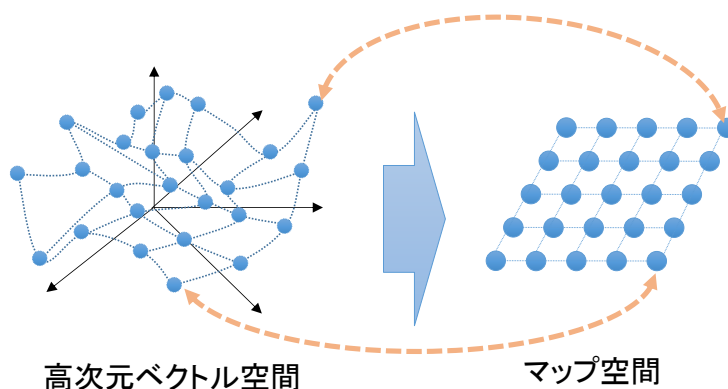


図 2-14 自己組織化マップ

入力データとノードデータの差を計算し、学習関数 h に基づいて、学習の度合いを決定する。これはマップ空間中における BMU とそれぞれのノードの間の距離に近い程、学習関数は大きな値をとり、距離が遠くなる程小さな値をとることから、入力データの影響を受けにくいことを意味する。

手順 2,3 を全入力データで行い、決めた学習回数分実行することによって入力データを反映したパターンデータの生成を行う。

2.8 時系列解析手法

一定時間間隔で測定され、時間とともに変化する時系列データである。各観測点での値に関係があるという仮定の下に分析・モデル化を行うことを時系列分析という。本研究で扱う電力使用量のデータは時系列データであり、この時系列分析を行うために時系列解析手法を用いる。

時系列データは統計的な性質が時間によらず変化しないものを定常時系列と呼び、変化するものを非定常時系列と呼ぶ。始めに定常系である自己回帰モデル、移動平均モデル、自己回帰移動平均モデルをそれぞれ説明し、その後非定常系である自己回帰和分移動平均モデル、季節自己回帰和分移動平均モデルについても説明する。

- 自己回帰モデル (AR: Autoregressive)

$$y_t = \alpha_0 + \sum_{i=1}^p a_i y_{t-i} + \varepsilon_t \quad (2.6)$$

ここで y_t は時刻 t の観測値、 ε_t は期待値 0 で散が一定であるホワイトノイズ、 a を自己回帰係数と呼ぶ。それぞれのデータはそれ以前の線形結合にノイズ（誤差）が加わって観測されることを意味する。

- MA (Moving Average model : 移動平均モデル)

1 次の AR は

$$\begin{aligned}
 y_t &= \varepsilon_t + a_1 y_{t-1} \\
 &= \varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 y_{t-2} \\
 &= \varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2} + a_1^3 y_{t-3} \\
 &= \varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2} + a_1^3 \varepsilon_{t-3} + \cdots \quad (2.7)
 \end{aligned}$$

と式(2.7)のように過去のホワイトノイズ(誤差)の線形和で表すことが出来る. つまり各データは過去の誤差に影響されるとしたモデルである.

この式変換を q 次まで行ったものを, q 次の移動平均モデル呼び以下のように表す.

$$\begin{aligned}
 y_t &= \varepsilon_t - b_1 \varepsilon_{t-1} - b_2 \varepsilon_{t-2} - \cdots - b_q \varepsilon_{t-q} \\
 &= \varepsilon_t - \sum_{i=1}^q b_i \varepsilon_{t-i} \quad (2.8)
 \end{aligned}$$

ここで b は移動平均係数である.

- ARMA(自己回帰移動平均モデル)

AR と MA を組み合わせたモデルを ARMA と呼ぶ. p 次の自己回帰過程において, 残差 ε_t が q 次の移動平均過程であるとき, 式(2-6), (2-8)を合わせて式(2.9)と表現される.

$$y_t = \sum_{i=1}^p a_i y_{t-i} - \sum_{i=1}^q b_i \varepsilon_{t-i} + \varepsilon_t \quad (2.9)$$

次数 (p, q) の ARMA と呼ぶ. 自己回帰も出ると移動平均モデルを結び付ければ簡潔に時系列データの特性を表現できるという考えに基づくものである.

- ARIMA(Autoregressive Integrated Moving Average model : 自己回帰和分移動平均モデル)

ARMA は定常時系列データを対象としており, 非定常データにトレンド性(時間によって統計的性質が変化)がある時系列データは非定常時系列データであるが, 各系列の階差に注目すれば定常的であると見なせる場合がある

$y_{n-1} = Ly_n$ となるラグ演算子 L を用いると一次の階差を x_t とすると

$$x_t = y_n - y_{n-1} = y_n - Ly_n = (1 - L)y_n \quad (2.10)$$

同様に 2 次の階差を z_t とすると, z_t は式(2.11)と表される.

$$\begin{aligned}
 z_t &= x_t - x_{t-1} = (1 - L)y_n - (1 - L)y_{n-1} \\
 &= (1 - L)y_n - (1 - L)Ly_n = (1 - L)^2 y_n \quad (2.11)
 \end{aligned}$$

d 次の階差を取った場合の ARMA は下式と表される.

$$\begin{aligned} a(L)(1-L)^d y_n &= b(L)\mu_n \\ a(L) &= (1 - a_1 L - a_2 L^2 - \dots - a_p L^p) \\ b(L) &= (1 - b_1 L - b_2 L^2 - \dots - b_q L^q) \end{aligned} \quad (2.12)$$

d 次の階差の時系列, p 次の AR モデル, q 次の MA モデルに適応したものを $\text{ARIMA}(p, d, q)$ という.

- SARIMA (Seasonal ARIMA : 季節自己回帰和分移動平均モデル)

時間単位に関する明らかな周期的変動 (例えば, 月次データであれば 12 ヶ月を 1 サイクルとする変動) が見られる場合, 非定常時系列データである, このような周期を季節性と呼ぶ. 12 カ月が周期であった場合, 月ごとの平均は異なるが, 12 カ月前の値から変化分を求めれば, 平均が 0 で定常に近い可能性がある. 季節性を考慮した ARIMA モデルのことを SARIMA という.

周期性が s である時系列データにおいて, データの季節階差を \tilde{y}_t とすると, 式(2.12)となる.

$$\tilde{y}_t = y_t - y_{t-s} \quad (2.12)$$

SARIMA は季節階差 $\{\tilde{y}_t\}$ をモデリングすることを考える. \tilde{y}_t が $s = 12$, つまり $12P$ (P は整数) 期前のデータに依存する ARMA モデルと考えると

$$\tilde{y}_t = \sum_{i=1}^P \Phi_i \tilde{y}_{(t-s*i)} - \sum_{i=1}^Q \theta_i u_{(t-s*i)} + u_t \quad (2.13)$$

ここで u_t は誤差項, Φ , θ はパラメータである.

近い時間帯でのデータとの相関を考慮するために, 誤差項 u_t が従う時系列モデルを考えると

$$u_t - \varphi_1 u_{t-1} - \dots - \varphi_p u_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad (2.14)$$

ε_t はホワイトノイズ, φ , θ はそれぞれパラメータである. 式(2.13),(2.14)で定義されるモデルが SARIMA である.

第3章 前提条件と提案電力データ共有手法

3.1 概要

本研究は家庭の電力データが各家庭から契約している各企業へ送られている状況を想定している。図 3-1 は電力提供企業が 3 社(A 社, B 社, C 社)の場合を表している。各企業はデマンドレスポンスなどのアプリケーションのために、自社のデータだけでなく、他社が保持するデータを利用したいというニーズがあるものとする。しかしながら、家庭の電力データは行動が推定可能であるため個人情報の一つであり、これらの企業は保持するデータを居住者の許可なく外部(他の企業)に実データを渡すことが出来ない。

このような状況下において、データを共有する手法は、匿名化、暗号化、パターン化が考えられる。まず、それぞれの企業が独自に匿名化(k-匿名化)を行うと、データ数が少ない場合にデータの損失が大きくなる。また、暗号に関する手法は、データを暗号化したまま計算処理を行う準同型暗号を用いたものを適応した手法が提案されているが、データを用いて複雑な計算を行うと計算コストが高くなり、処理できるサーバーなどの制約が発生する恐れがある。パターン化は代表的なパターンを電力データの場合はパターン分類するための教師データとなるものがない。そして各企業がパターン化を行ったデータを共有する手法を用いると各企業が保持するデータの傾向が流出してしまい、契約者数が少ない企業に関しては個人を特定されてしまう恐れがある。

以上データ共有に関する問題点があるが、本研究ではパターン化に着目している。代表パターンのラベルが存在しないことを上述したが、教師なし学習の SOM を用いることで入力データの分布を考慮して代表パターンを作成する。SOM の学習アルゴリズムは、マップ空間中のノードデータと入力データの比較により、データを学習するので、入力データ対入力データを比較することがない。つまり、各企業が保持するデータやパターンデータを互いに比較せず学習をする。

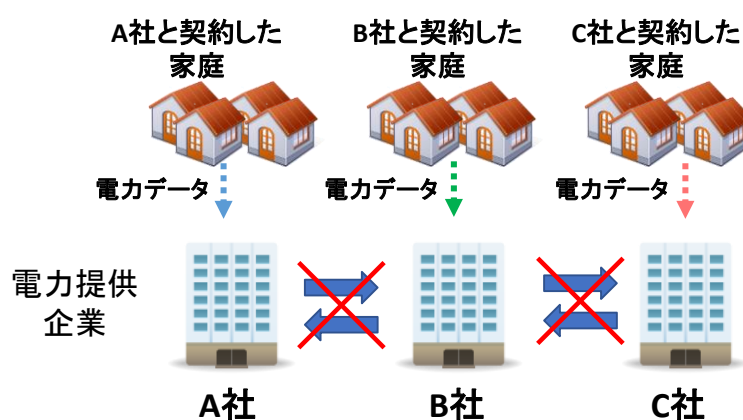


図 3-1 想定する前提条件

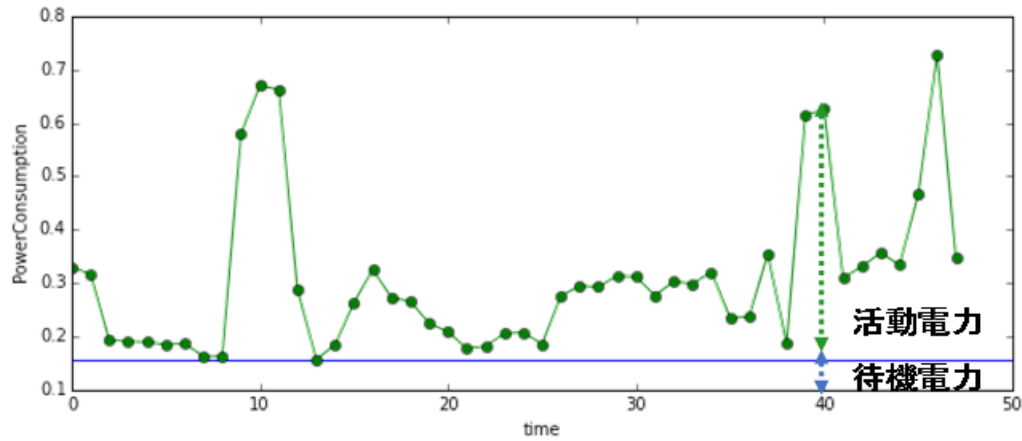


図 3-2 待機電力と活動電力の分離

これによりお互いに持つ実データやパターンデータを直接公開することがなく、プライバシーの問題を解決する. この性質を用いて本研究では SOM を用いた電力データ共有手法を提案する.

3.2 提案する SOM を用いた電力データ共有手法

3.2.1 電力データ

本研究で扱う電力データは, 1 軒分を 30 分間隔で送信するスマートメータの値を 1 つのデータと想定している. i 番の家庭の電力データ x_i は, N_{home} を全体の家庭総数とすると, 以下のよう

$$x_i = (x_{i,00:00}, x_{i,00:30}, \dots, x_{i,23:30}) \quad i = 1, 2, \dots, N_{home} \quad (3.1)$$

本研究ではパターン化したデータを電力ピークシフトのためのデマンドレスポンスへ適応することを目標としており, 電力データをパターン化するにはデマンドレスポンスへの適応可能性を考慮する必要がある. 家庭の電力は待機電力と人間の活動により発生する活動電力の二つに大きく分けることができる. 待機電力は固定的なものであり, デマンドのピークシフトが可能な部分は活動電力である. 本研究ではこの活動電力部分を入力データとして取り扱う. 活動電力は総電力使用量から待機電力を引いたものとし, 待機電力は各家庭の電力使用量の最小値とした(図 3-2).

3.2.2 SOM を用いたデータ共有アルゴリズム

本提案手法はデータの学習によるパターンデータ生成と, カウントの 2 つのステップに分かれている. まず, 各企業が保持するデータのパターン化を行う. そのパターンデータを新たな入力データとして全体のデータを考慮したパターンデータを生成し, そのパターンデータを共有する. 次に, 各企業は保持するデータとパターンデータの類似度を計算し, それぞれのパターンデータに近い家庭が何軒存在するかをカウントする. その後カウント情報を全体で足し合わせ, 全体としてどのパターンの家庭がそれぞれ何軒存在するかという情報を生成する. 以上の処理を行うことで安全に各企業が保持するデータの傾向を考慮したパターン生成とそのパターンデータと

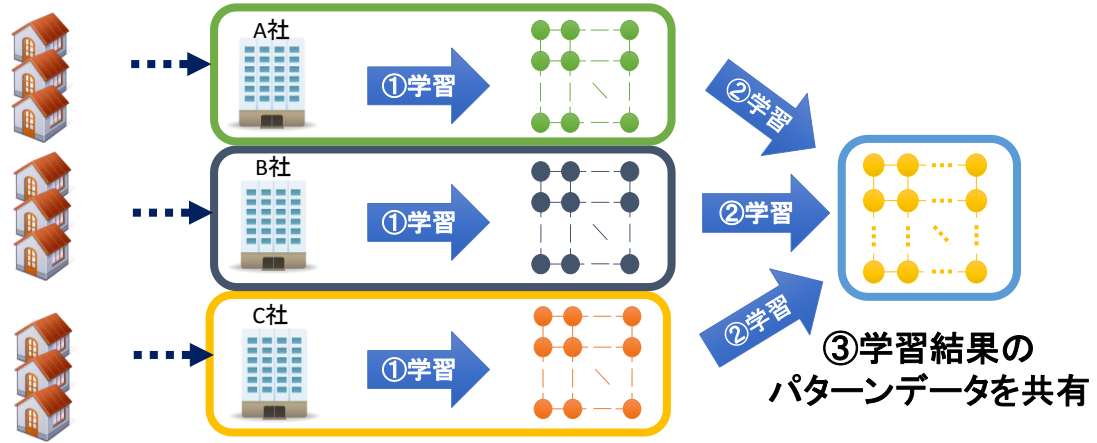


図 3-3 学習ステップ

類似度が高いと判断されたデータ数の数え上げを行う。二つのステップについて詳細を次に述べる。

- データの学習によるパターンデータ群生成

各企業がそれぞれ保持するデータの傾向を学習するために本手法では 2.7 節で解説した SOM を用いる。SOM のマップ中のノードがパターンデータに対応する。学習ステップにおける全体図を図 3-2 に示す。学習ステップではまず各企業がそれぞれ保持するデータを用いて独立でパターン生成を行う。SOM のマップ中の学習回数 t におけるノードベクトルを $\mathbf{w}_i(t)$ とする。パターン生成は以下の SOM のアルゴリズムで行う。

1. マップ空間中に初期ノードを設定

SOM のマップ空間に初期ノード($\mathbf{w}_i(0) \ i = 1, 2, \dots, N_{map}$)を乱数で生成する。 N_{map} は SOM のマップ空間における全ノード数である。

2. 入力データ(各家庭の電力消費量一日分)に最も類似度が高いノードの選択

類似度の計算はユークリッド距離に基づいて行う。最も類似度が高いノードを入力データに対する BMU とする。式(3.2)と表される。ここで、 x は入力となる各家庭の電力データ(48 次のベクトル)である。

$$BMU = \underset{i}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{w}_i(t)\| \ i = 0, 1, \dots, N_{map} \quad (3.2)$$

3. マップ空間中のノードの持つ値の更新

下式を用いて BMU の周辺ノードを入力ベクトルに近づける。

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{BMU,i}(t)(\mathbf{x} - \mathbf{w}_i(t)) \quad (3.3)$$

$$h_{BMU,i}(t) = \alpha(t) \exp\left(-\frac{d_{BMU,i}^2}{2\sigma(t)^2}\right) \quad (3.4)$$

$$\alpha(t) = \left(1 - \frac{t}{T}\right) \quad (3.5)$$

第 3 章 前提条件と提案電力データ共有手法

ここで、 T は全学習回数、 t は現在の学習回数である。次に各企業で生成されたパターンデータを新たな入力データとし、上述した SOM のアルゴリズムを用いてデータを学習する。

これらの 2 段階の学習によって出力されるパターン群のデータは各企業が保持するデータを考慮した代表パターンとして各企業に共有し、学習ステップを終了する。

$$w_i = (w_{i,00:00}, w_{i,00:30}, \dots, w_{i,23:30}) \quad i = 1, 2, \dots, N_{map} \quad (3.6)$$

● データ数カウントステップ

データ数カウントステップは、それぞれのパターンデータに近い家庭が全体で何軒存在するかという情報を取得するためのステップである。まず、各企業は学習ステップで生成されたパターンデータと保持するデータの類似度から代表パターンを選択する。類似度が高いデータの選択は SOM の学習アルゴリズム中の BMU の探索と同じ式(3.2)、つまりユークリッド距離が最も小さいパターンのデータ(ノードデータ)を選択する。選択されたノードはカウントを 1 追加し、これをすべて企業の保持するデータに対してそれぞれの企業が独立で行う。

次に、それぞれの企業が生成したパターンデータ全体に対するカウントの足し合わせを行う。足し合わせにおいて 3 ケースを想定する。「カウント情報は秘匿にすべき情報で企業が結託する可能性がある」、「カウント情報は秘匿にすべき情報で企業が結託しない」、「カウント情報は秘匿にしないでよい」。それぞれの場合について以下で述べる。ここで A 社、B 社、C 社のカウント情報 a , b , c とする。

(ア) カウント情報は秘匿にすべき情報で、企業が結託する可能性がある

それぞれのカウントの足し合わせを秘匿に行うために、安全度が高い暗号化したままの計算を行うことが可能である準同型暗号を用いる。全体図を図 3-4 に示す。ここでは加算計算のみを行うため、Paillier や Additive ElGamal といった準同型暗号の中でも計算コストが比較的少ない暗号を用いる。公開鍵を各企業に配布し、その公開鍵を用いて暗号処理を各企業で行う。それぞれの暗号化結果を $Enc(a)$, $Enc(b)$, $Enc(c)$ として送信し、それぞれの企業の結果を暗号化したまま加法演算 $Enc(a + b + c)$ を行う。最後に秘密鍵を用いて復号することによって全体のカウント合計値を得ることが可能となる。

(イ) カウント情報は秘匿にすべき情報で、企業が結託しない

企業が互いに結託しないという前提が成立する場合は図 3-5 に示した巡回型のカウントをすることで自社のカウント情報を公開せずに全体の合計値を計算することが可能である。

初め A 社がカウント数を乱数(r)で決める。この乱数情報は A 社のみで管理する。次に A 社のカウント情報(a)を先ほど生成した乱数のカウント(r)に加算する。この結果($a + r$)を B 社に送信し、B 社も同様に自社のカウント情報を受け取った結果に加算する($a + b + r$)。同様に C 社でもカウントを追加し($a + b + c + r$)する。すべての企業のデータが加算されたら、最初にデータを反映した A 社にデータに送信する。A 社は加算結果から最初に生成した乱数のカウントを引く($a + b + c$)。この結果が全体のカウント情報となる。前提として企業が結託しないことを上げている理由は、例えば B 社と C 社が協力して A 社のカウント情報を計算することが可能であることを考慮するためである(C 社は最初に A が生成する乱数 r が分かるので、B 社はこの乱数 r の値を用いることで A 社のカウント a を計算が可能である)。この手法は(ア)の場合と異なり暗号化

第3章 前提条件と提案電力データ共有手法

コストがなくなるため、計算量が少ないというメリットがある。

(ウ) カウント情報は秘匿にしなくてもよい

各企業がパターンデータのカウント情報を出すことによる侵害の危険がないとする場合には特別な処理は必要がなく、それぞれの実カウント情報を一つのデータテーブルに収集する。

以上のカウントの合算を終了すると、地域全体としてそれぞれのパターンの家が何軒あるという情報を生成することが出来るため、デマンドレスポンスの計画や、その他アプリケーションへの適応をすることが可能である。

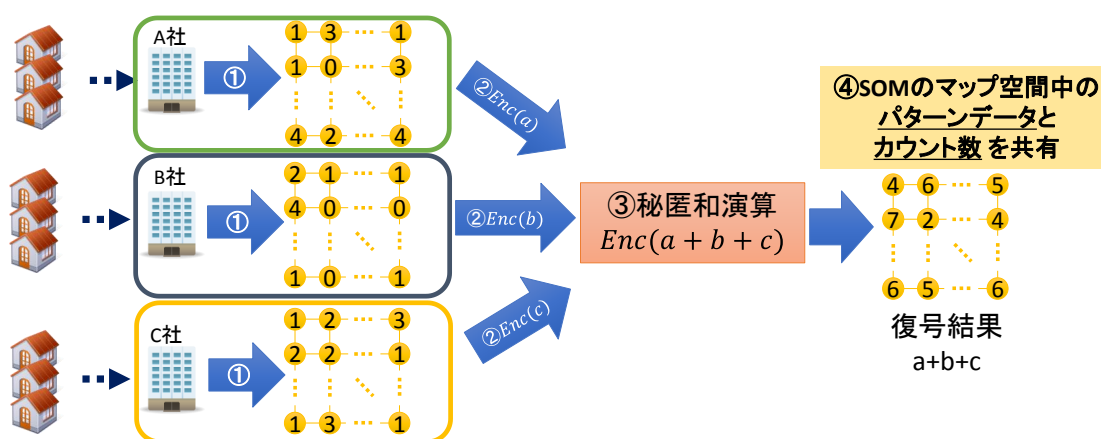


図 3-4 カウント情報は秘匿にすべき情報で、企業が結託する可能性がある場合のカウント情報集計

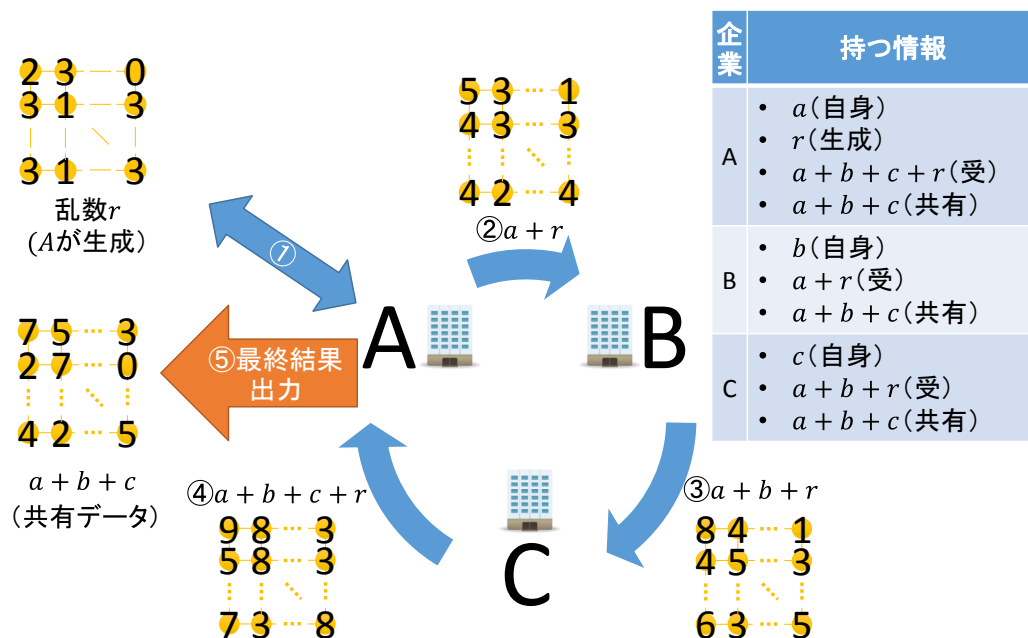


図 3-5 カウント情報は秘匿にすべき情報で、企業が結託しない場合のカウント情報集計

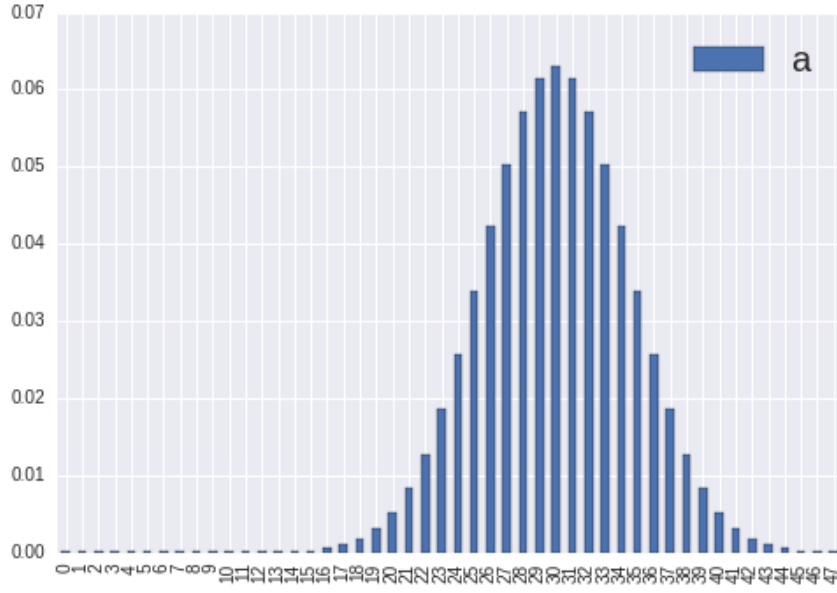


図 3-6 ピーク時間が 18 時の場合の重み付け分布

3.2.3 デマンドレスポンスを目的とした重み付け

本研究のデータ共有後のデータ利用想定アプリケーションは地域のデマンドレスポンスである。デマンドレスポンスの計画を行うために、ピーク時間における実データとノードデータ(パターンデータ)類似度を考慮したデータの精度が必要であるが、通常の BMU の選択は、ユークリッド距離が最も小さくなるものとしており、ピーク時間の類似度が低くなる可能性がある。ピーク時間におけるデータの精度を高めるために、再近傍ノードの探索時に重み付けを行う。ここでは電力需要予測を行うことでピーク時間の推定を行う。重みづけはピーク時間を中心とした正規分布に従う重みづけパラメータ a_t を設定する。電力需要のピーク時間が 15 時であった場合の正規分布の重みづけは図 3-6 のようになる、これを踏まえた BMU の式(3.2)は以下ようになる。

$$BMU = \underset{i}{\operatorname{argmin}} \left(\sqrt{\sum_{t=1}^T a_t (x_t - w_{i,t})^2} \right) \quad (3.7)$$

$$a \sim \mathcal{N}(p, \sigma^2) : a(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-p)^2}{2\sigma^2}} \quad (3.8)$$

ここで T は次元数(スマートメータの場合 48)、 x を電力データ、 w が SOM ノードデータ(パターンデータ)、 a は重み変数 p は予測需要ピーク時刻、 σ^2 は正規分布分散値(重みづけ調節変数)である。重みづけの強度は正規分布の分散値 σ^2 に依存し、 σ^2 が小さい値を取るほど、ピーク時間のデータの類似性の重みづけが強くなる。

3.3 翌日需要電力のピーク時刻推定

3.2.3 項で用いる重みづけは地域の需用電力ピーク時間がパラメータの一つとなっている。ピーク時間の予測を行うために、本研究では翌日需要電力の推定を行い、その推定値からピーク時間を推定する。電力需要の予測手法は電力データと気温情報を組み合わせて行う重回帰分析を用いたものや、サポートベクターマシンを用いたものなど様々な研究が行われているが、本研究ではスマートメータのデータのみで行うことが可能である、時系列解析手法の2章で述べた SARIMA (季節自己回帰和平均モデル) を用いている。SARIMA は ARIMA モデルに季節項を組み合わせてものである。家庭の電力データは一日などの周期性が見られるため有効であると考え利用した。以下に需要予測について説明する。

本研究の前提として同じ地域の家庭が違う企業と契約していることを想定している(図 3-8)。まずそれぞれの企業は契約している家庭のデータの合計値を計算する。その合計値のみをすべての企業で統合(足し合わせ)することによって個々の家庭のデータを出すことなく地域需用電力の計算を行う。

全体の合計値を計算後この値を地域の合計消費電力として SARIMA モデルを構築する。モデル構築には入力を7日間(48×7 samples)として1日(48 samples)先予測を行う。この予測結果に基づきピーク時刻を推定するものとした。

SARIMA は $(p, d, q) \times (P, D, Q)_s$ で表されるが、自己相関、偏自己相関の値からモデルを絞り、それらのモデルを構築し、モデルの評価値を AIC (赤池情報量基準) として、それぞれ最終的な時数 $(p, d, q) \times (P, D, Q)$ を決定する。手順を以下に示す。

- 周期性の決定(s)

スマートメータデータ 30 分値を用いていることから、 $s=48$ が周期である。

- 階差項 d と D を単位根検定により決定

定常性を統計的に判断するために、単位根検定として拡張ディッキーフラー (ADF) 検定を行う。

- 単位根検定により決定した d, D を用いて、 p, q, P, Q を複数の組み合わせでモデルを作成

自己相関、偏自己相関の値からモデルを絞る。図 3-7 の場合においては $p=4, q=1$ 付近のパラメータとなることが推定出来る。(s=48 も確認される)



図 3-8 翌日需要電力のピーク時刻推定の流れ

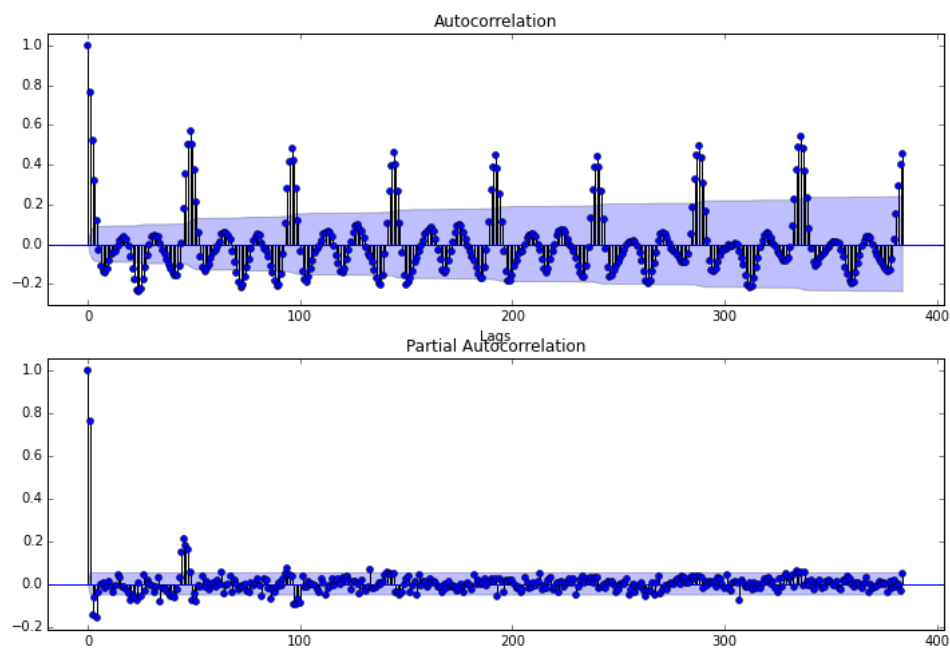


図 3-7 自己相関(ACF)と偏自己相関(PACF)

- 作成したモデルを用いて AIC が最小となるものを探索し、最適な p, q, P, Q を決定する

モデル構築は、説明変数が少ない状態で当てはまり能力が高くなる説明変数の数を決定する必要がある。適切なパラメータ数を決めるために AIC((Akaike's information criterion))を用いた。

$$AIC = \ln\left(\frac{S_e^2}{n}\right) + \frac{2K}{n} \quad (3.9)$$

ここで S_e は残差, n はサンプル数, K は説明変数の数である。SARIMA における、モデル化と予測イメージ図 3-9, 予測の例を図 3-10 に示す。

分析(入力) 予測(出力)

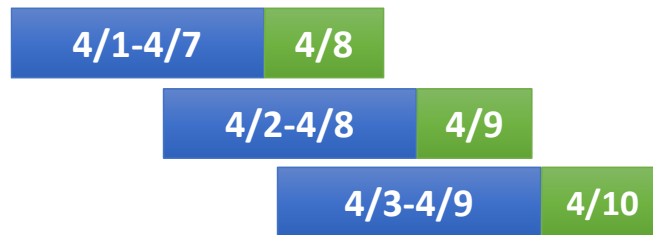


図 3-9 SARIMA における, モデル化と予測イメージ

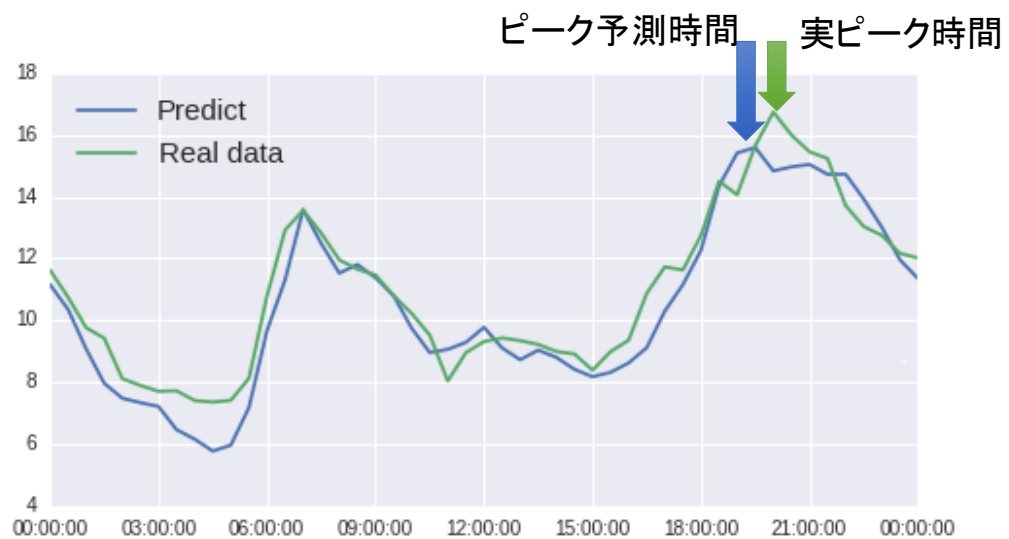


図 3-10 SARIMA を用いた一日先需要予測結果と実データ

第4章 データ取得環境の構築

本論文で提案する手法の評価に用いる電力データは、2 章で紹介した川崎市スマートシティプロジェクトで得られたデータを用いている。本章ではスマートシティプロジェクトのシステムについて述べる。

4.1 川崎市スマートシティプロジェクトにおけるシステム

川崎市スマートシティプロジェクトにおいては、スマートグリッドの技術標準である IEEE1888 を用いたシステムで稼働している。全体図を図 4-1 に示す。本プロジェクトの対象としている家庭は武蔵小杉駅周辺の家庭約 150 件で、取得可能なデータは築年数や延べ床面積などの世帯情報、子ブレーカー単位の消費電力やエネファーム発電量などの電力利用情報、年齢や性別などの住民情報である。プロジェクトに参加している家庭の中でも特にデータの第三者提供を許可した家庭のデータのみを本研究では用いている。

4.2 IEEE1888

4.2.1 概要

エネルギー消費量の見える化、無駄や改善の診断、快適自動制御、電力需給バランス制御といったエネルギー管理システムは ICT 技術で担おうとしており、この動きのことを一般にグリーン ICT と称されている。しかしながら BEMS や HEMS 等の設備は、独自のデータベースを持った形で、各社で開発されており、他社製品との連携は困難である。また、従来よりも広い単位で

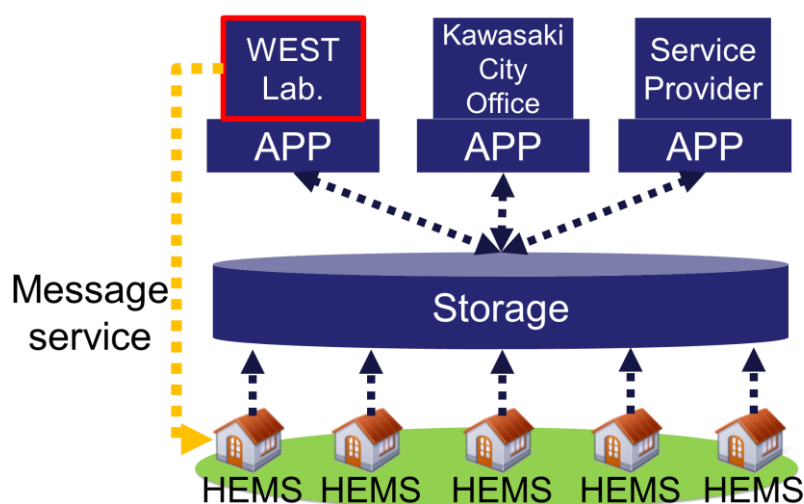


図 4-1 川崎市スマートシティプロジェクトにおけるシステム全体図

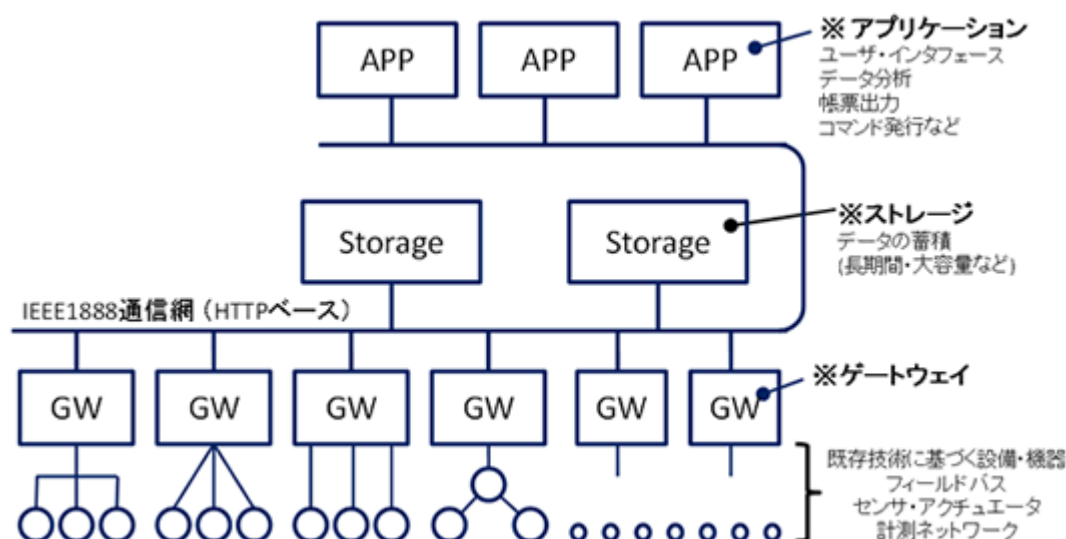


図 4-2 IEEE1888 で構成されるシステム [21]

エネルギー管理が行われようとしているが統一的な手法が無いため、データ流通の利便性を欠くシステム形態となっている。

通信プロトコルである IEEE1888 は、このような背景から、従来の設備管理の枠を超え、エネルギー管理を中心に添えた運用に視点を広げ、グリーン ICT 技術の社会インフラ化を目指して設計・開発された。データ蓄積、マルチベンダー化、そして広域管理を可能にするアーキテクチャが採用され、これまで設備管理システムの導入が困難であった中小規模施設や家庭などにまで、技術インフラとして浸透できるように設計されている。IEEE1888 により、小規模施設の設備情報が、オンライン化され、広域にわたる情報の集中監視、見える化、遠隔制御が実現可能になる。

具体的には、図 4-2 のように [21]、各センサーや設備の情報を送受信する各フィールドバスからのデータを、ゲートウェイ経由でデータベースなどのストレージや、ウェブブラウザなどのアプリケーションと相互接続して連携させる。これらのデータから、リアルタイムでの設備情報、過去の計測データの閲覧、それらのデータに基づいた制御が可能となる。IEEE1888 は商業施設や事務所施設の通信プロトコルを対象としており、外部との通信を想定して設計されている。

IEEE1888 で構築されるシステムは、ゲートウェイ、ストレージ、そしてアプリケーションと呼ばれる機能部品で構成され、これらの部品間は HTTP ベースの通信で情報が交換される。IEEE1888 はこの通信の手順やフォーマットを規定したものであり、それぞれの機能部品は次のような役割を果たす。

- ゲートウェイ

センサーやアクチュエータが搭載されている他のシステムに接続し、取得したデータを IEEE1888 で扱えるようにする。または IEEE1888 で送られてきたコマンドを発行する。ゲートウェイは必ず TCP/IP ネットワークに接続され、IEEE1888 の通信インターフェースを提供する。接続の対象となるシステムには、各ベンダが独自に規定している通信バスや、既存の Lonworks、

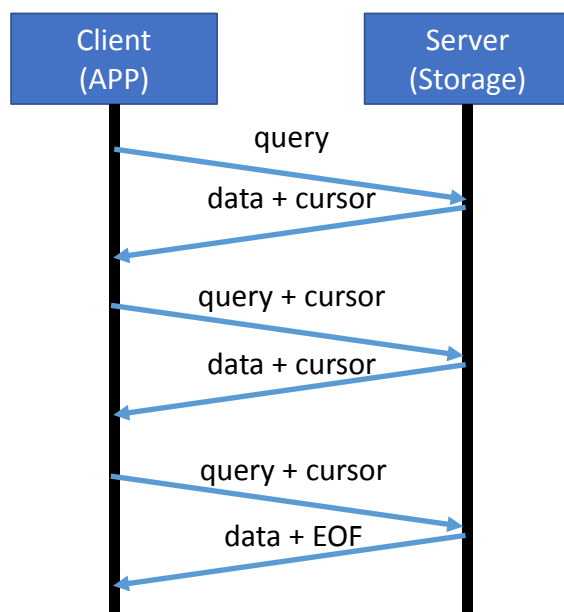


図 4-3 Fetch 手順(通信シーケンス)

BACnet, ZigBee, RS485, 1-Wire, PCL などのローカル通信バスも含まれる。

- ストレージ

IEEE1888 フォーマットで表現されたデータを長期間にわたって蓄積し、様々なアプリケーションとのデータ共有プラットフォームとして使われる。データ分析のような計算処理は、何らかのアプリケーションと連携させることによって達成される。

- アプリケーション

ストレージからデータを読み出し、運用管理者とのインターフェースとなる。ゲートウェイに対してコマンドを発行するのもアプリケーションの役割である。また、データ分析を行うアプリケーションは、ストレージからデータを読みだした後、分析処理を行い、再度ストレージに書き込むといった動作をする。このように、IEEE1888 はデータの蓄積を中心に据えたアーキテクチャを採用している。

川崎市スマートシティプロジェクトのシステムは、この IEEE1888 を用いて構築を行った。IEEE1888 のデータの取得とデータのフォーマットの詳細について述べる。

4.2.2 データの取得

本システムにおけるデータの Fetch(取得)に関する説明をする(図 4-3)。データを提供する側を Server、取得の要求をする側を Client とする。まず、Client は要求するデータを記述したメッセージ(query)を発行し、Server 側に送信することにより Fetch 手順が始まる。query のメッセージには type に関する記述をするが Fetch では 'type=storage' とする。query を受け取ったサーバーは query に該当するデータを検索し、Client 側へデータを送信するが、ここで送信される

```

<body>
  <pointSet id=http://*.co.jp/service_provider_common/home_no_1>
    <pointSet id=http://*.co.jp/service_provider_common/home_no_1/hems_no_1/>
      <point id=http://*.co.jp/service_provider_common/home_no_1/
        hems_no_1/measure_point_id_10001>
        <value time="2015-05-15T00:00:00.000+09:00"> 0.249</value>
        <value time="2015-05-15T00:30:00.000+09:00"> 0.235</value>
        <value time="2015-05-15T00:30:00.000+09:00"> 0.211</value>
      </point id>
    </pointSet>
  </body>

```

図 4-4 body クラスの表現

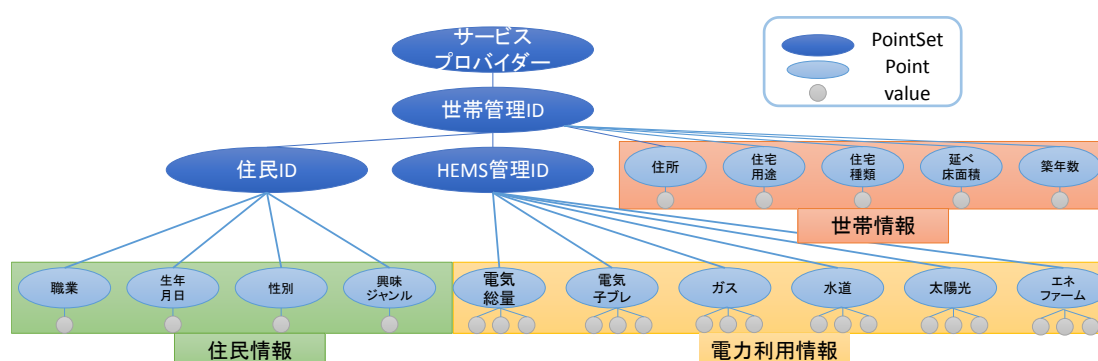


図 4-5 全世帯の情報取得時のポイントツリー

データ量は `acceptableSize` で定義したデータ量を超えないようにする。送信すべきデータが残っている場合は `cursor` 情報として応答するデータと共に `Client` に送信する。`Client` は受け取ったデータを確認し、`cursor` が付随していたら最初に送信した `query` を `cursor` とともに再送する。`Server` は `cursor` が持つ情報から前回送信をしていなかったデータを送信する。ここで送信すべきデータが `acceptableSize` を再度超えてしまった場合は先ほどと同様に、`cursor` を付随してデータを送信するが、`Client` が発行した `query` に当てはまる条件を満たしたデータをすべて送信することができる場合は `cursor` でなく、`EOF` として `Client` へ返す。このシーケンスにより `Fetch` が完了する。この `Fetch` のプログラムを `PHP` で実装した。

4.2.3 データのフォーマット

前項でデータ取得のための通信シーケンスを示したが、ここでは `IEEE1888` のデータの詳細について述べる。`IEEE1888` で交換されるメッセージは `transport` と呼ばれ、`header` と `body` で構成されている。`header` にはクエリや処理の成否が格納され、`body` に計測値やエラー情報などが格納される。

`header` 中の `query` クラスは問い合わせ内容を示すもので、その中には対象とするデータ範囲を示す `key` オブジェクトが複数ある。`key` は判定用プロパティ、`eq(equal)`や、`gt(greater than)`などを持ち、要求するデータ範囲を表す。

第 4 章 データ取得環境の構築

body クラスは複数の pointSet オブジェクトあるいは複数の point オブジェクトを持つ。pointSet は複数の pointSet オブジェクトあるいは point オブジェクトをもつ。Point クラスは id プロパティによってポイント ID を指定し、複数の value オブジェクトを持つ。value クラスは時系列データの要素を表しており、time 属性に指定された時間を表現する。具体的な例を図 4-4 に示す。例では pointSet id として home_no_1 を選び、さらにその下の pointSet では hems_no_1 を選んでいる。さらにその下にある point id (measure_point_id_10001) は value を表している。

pointSet, point, value は上述した階層構造をとっている。本プロジェクトにおける階層構造を図 4-5 に示す。データは大きく世帯情報、電力利用情報、住民情報に分かれるが、世帯情報と住民情報は時間によらず一意であるため point と value は一対一の関係となっている。一方で電力利用情報は時系列データであるため、point の下に複数の value がある関係となっている。そのため電力利用情報を取得する際には key の判定用プロパティでデータ範囲(時間範囲や value の範囲)を指定することで目的のデータを取得する。

第5章 評価

5.1 動作環境, 対象データ

本研究に用いるデータは第3章で紹介した川崎市武蔵小杉プロジェクトのデータを用いて評価を行う。欠損値などを考慮し、本研究の評価には57軒分のデータを用いた。より大きい規模を想定した場合を考えるため commission for regulation によって提供されている Irish smart meter dataset [22] を用いた。この中のデータ1000軒分データを評価対象として用いた。二つのデータセットはどちらもスマートメータの電力量で30分値である。これらの分析にはPythonを利用した。

5.2 k-匿名化による情報損失度の比較


2章で紹介した情報保護指標である、k-匿名化による情報損失度の比較を行う。k-匿名化とは同じ属性を持つ人がk個以上データテーブルに存在する状態のことを表す。本研究では家庭の電力データを対象としており、k匿名化処理を施すと図5-1のようになる。これはHome1-3とHome4-6を二つのクラスタに分けたとき、クラスタ内のデータの平均値(あるいは最大値最小値)などの代表値のみで表現することで匿名化を行う。

k-匿名性を満たすため本研究では2章で示したk-member clusteringを用いた。この手法では始めにランダムに初期データをデータテーブルの中から選択し、そのデータの値に対して類似度が高い順にさらにデータを選択していき、サイズがk個となった際にそのデータ集合を一つのクラスタとする。次から新たなクラスタを生成し、これをすべてのデータが一つのクラスタにラベリングされるまで続ける。クラスタリング処理後、そのクラスタの代表ベクトルとして、クラスタに含まれるデータの平均ベクトル(重心)を採用する。情報損失度はその代表ベクトル(c_i)とクラスタ内の各データ値(x_i)との差によって評価する。

k-member clustering を用いた匿名化処理を「提案手法を用いて全体データを把握して匿名化処理を行った場合」と、「企業が互いに共有せずそれぞれ独立に匿名化処理を行った場合」で比較を行う。企業数は一つの変数として扱い(n_p)、各企業に分配される数が変化するものとした。各企業が持つデータ数は、全データ/企業数としてデータを分配して計算を行うものとする。(全体データサイズが $N_{home} = 1000$ で企業数が $n_p = 10$ であれば各企業のデータサイズは100)提案手法を用いて匿名化を行った場合のMAEと提案手法を用いずに独立的に匿名化を行った場合のそれぞれの誤差を $MAE_{n_p, proposed}$, $MAE_{n_p, Independent}$ とし、比較の指標としてRateを以下のように定義する。

$$MAE = \frac{1}{N_{home}} \sum_{i=1}^{N_{home}} |c_i - x_i| \quad (5.1)$$

	0:00	0:30	1:00
x_1 (Home1)	209	343	234
x_2 (Home2)	243	143	121
x_3 (Home3)	249	214	234
x_4 (Home4)	895	543	432
x_5 (Home5)	634	346	765
x_6 (Home6)	644	543	345



	0:00	0:30	1:00
c_1	234	233	196
c_2	234	233	196
c_3	234	233	196
c_4	724	477	514
c_5	724	477	514
c_6	724	477	514

図 5-1 電力データの k-匿名化

$$Rate_{n_p} = \frac{MAE_{n_p, Proposed}}{MAE_{n_p, Independent}} \quad (5.2)$$

つまり Rate が 1 より小さい場合、提案手法が有効であることを示す。二つのデータセットにおける評価を次に記す。

5.2.1 武蔵小杉データセットを用いた評価

武蔵小杉のデータを用いた評価を行った場合の結果を示す。SOM の MapSize を 5×5 と 10×10 の場合でそれぞれ計算を行った。MAE の計算結果を図 5-2、図 5-3、Rate の計算結果を表 5-1、表 5-2 に示す。k が大きく、企業数が増加するほど提案手法が優位であることを示している。特に Map Size が 10×10 、 $k = 8$ 、 $n_p = 4$ のときは $Rate = 0.88$ つまり k-匿名化を行う際の情報損失 12% 削減を達成した。しかし、k が小さい ($k = 2, 3$) の場合においては、独立的に匿名化を行った場合の方が良い結果となる場合が多い。この理由は、独立的に匿名化を行ったとしても $k=2, 3$ といった k が小さい場合では類似度の高いデータを発見することが、各企業の保持するデータ数が少ない状態でも可能であり、一方で提案手法では共有することでデータ数は多くなるが、パターン化を行った際の実データの誤差が大きくなるため情報損失も大きくなると考えられる。

5.2.2 Irish データセットを用いた評価

武蔵小杉データセットを用いた評価を行った場合と同様に Irish データセットを用いた場合の結果を示す。SOM の MapSize を 10×10 と 20×20 の場合でそれぞれ計算を行った。MAE の計算結果は図 5-2、図 5-3 のようになり、Rate の計算結果を表 5-1、表 5-2 示す。k が大きく、企業数が増加するほど提案手法が優位であることを示している。特に SOM のマップサイズが 20×20 で $k = 20$ 、 $n_p = 15$ のときは $Rate = 0.80$ 、つまり k-匿名化を行う際の情報損失を 20% 削減することが出来ている。しかしながら、k が小さい ($k = 2, 3$) の場合においては、独立的に匿名化を行った場合の方が良い結果を示した。この理由は武蔵小杉データセットと同様に提案手法ではパターン化を行った際の実データの誤差が大きくなるためであると考えられる。

また、Irish データセットの評価において、データサイズによる影響を評価するため家庭全体数 $N_{home} = 100, 500, 1000$ の場合についてそれぞれ評価を行った。SOM のマップサイズを 20×20 、 $n_p = 10$ と設定した。結果を図 5-6、表 5-5 に示す。k ≥ 6 のとき、同じ k に対してデータ

第5章 評価

数 N_{home} が大きいく程, MAE(誤差)が少なくなっている. これは k -匿名化のための k -member clustering 処理を行う際に類似度が高いデータを選択する場合, データ数の多い方がより類似度を選択することが可能であるためだと考えられる.

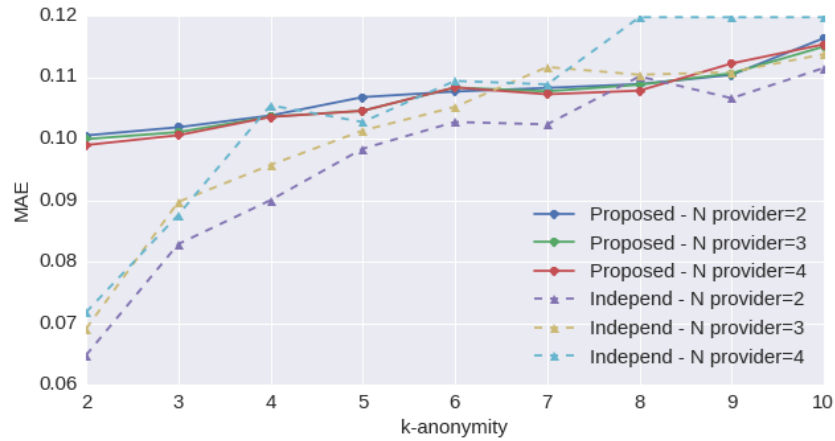


図 5-2 k と企業数の関係 MapSize 5×5 (武蔵小杉データセット)

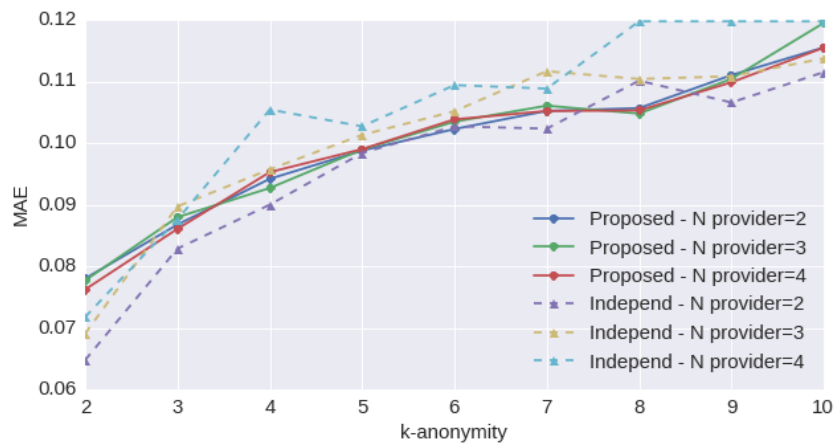


図 5-3 k と企業数の関係 MapSize 10×10 (武蔵小杉データセット)

表 5-1 MapSize 5×5 における $Rate_{n_p}$ (武蔵小杉データセット)

$n_p \backslash k$	2	3	4	5	6	7	8	9	10
2	1.55	1.23	1.15	1.09	1.05	1.06	0.99	1.04	1.04
3	1.45	1.13	1.08	1.03	1.03	0.96	0.99	1.00	1.01
4	1.38	1.15	0.98	1.02	0.99	0.99	0.90	0.94	0.96

表 5-2 Map Size 10×10 における $Rate_{n_p}$ (武蔵小杉データセット)

$n_p \backslash k$	2	3	4	5	6	7	8	9	10
2	1.20	1.05	1.05	1.00	1.00	1.03	0.96	1.04	1.04
3	1.13	0.98	0.97	0.98	0.98	0.95	0.95	1.00	1.05
4	1.06	0.98	0.90	0.96	0.95	0.97	0.88	0.92	0.96

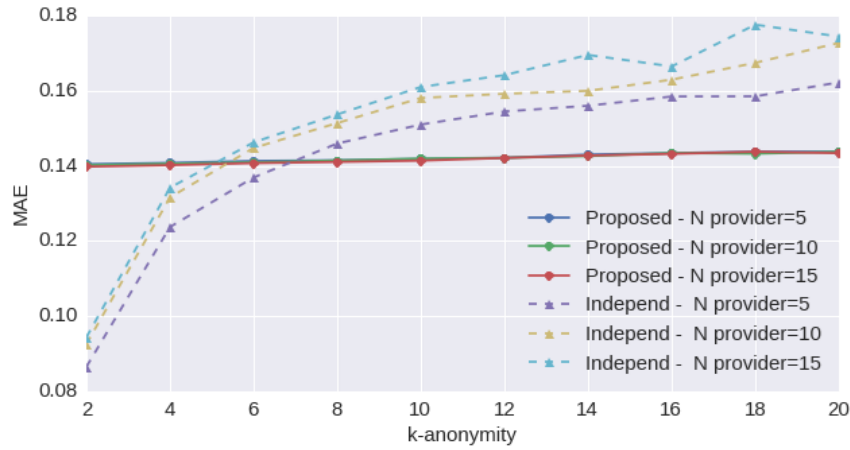


図 5-4 k と企業数の関係 MapSize 10×10 (Irish データセット)

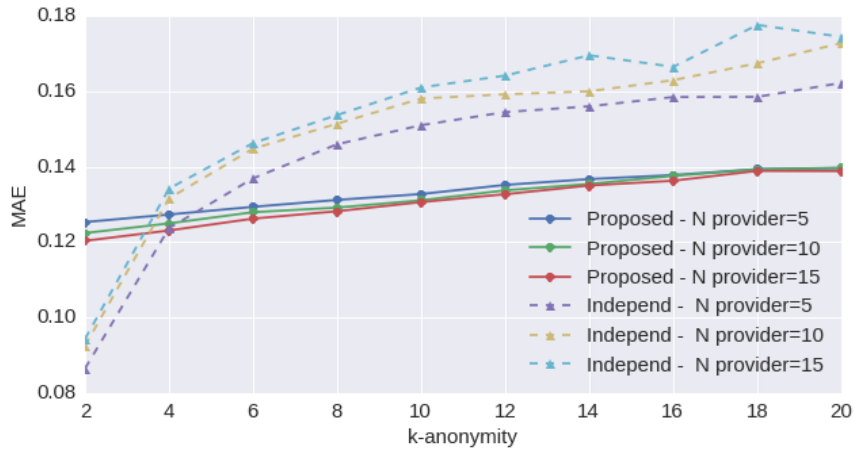


図 5-5 k と企業数の関係 MapSize 20×20 (Irish データセット)

表 5-3 MapSize 10×10 における $Rate_{n_p}$ (Irish データセット)

$n_p \backslash k$	2	4	6	8	10	12	14	16	18	20
5	1.62	1.14	1.03	0.97	0.94	0.92	0.92	0.90	0.91	0.89
10	1.51	1.07	0.97	0.93	0.90	0.89	0.89	0.88	0.86	0.83
15	1.48	1.04	0.96	0.92	0.88	0.87	0.84	0.86	0.81	0.82

表 5-4 MapSize 20×20 における $Rate_{n_p}$ (Irish データセット)

$n_p \backslash k$	2	4	6	8	10	12	14	16	18	20
5	1.45	1.03	0.94	0.90	0.88	0.88	0.88	0.87	0.88	0.86
10	1.32	0.95	0.88	0.85	0.83	0.84	0.85	0.84	0.83	0.81
15	1.28	0.92	0.86	0.83	0.81	0.81	0.80	0.82	0.78	0.80

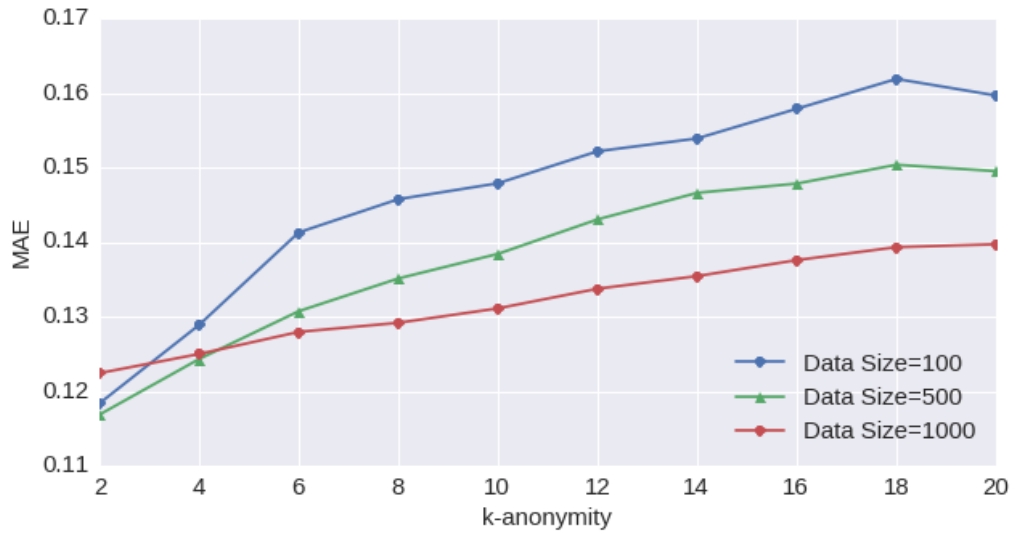


図 5-6 データサイズを変化させた場合の MAE による評価

表 5-5 データサイズを変化させた場合の MAE による評価

k	$N_{home} = 100$	$N_{home} = 500$	$N_{home} = 1000$
2	0.118	0.117	0.122
4	0.129	0.124	0.125
6	0.141	0.131	0.128
8	0.146	0.135	0.129
10	0.148	0.138	0.131
12	0.152	0.143	0.134
14	0.154	0.147	0.135
16	0.158	0.148	0.138
18	0.162	0.150	0.139
20	0.160	0.149	0.140

5.3 デマンドレスポンスを想定した評価

本研究では全体データを考慮したパターン化によるデータ共有手法を提案したが、このパターンデータを用いてデマンドレスポンス計画に向けた需用電力の計算を行う。提案手法ではパターンデータを生成し、それぞれの実データと最も類似度が高いものを選択し、その選択されたデータを実データの代わりに用いる。需用電力を計算するためにはそれぞれの家庭の電力需要量を合算するため、それぞれの家庭のデータの精度が必要である。本節では実データとパターンデータの差を計算する。

本研究では 3.2.3 項で提案したデマンドレスポンスを目的とした重み付けの設定するためのパラメータ σ^2 を変化させることにより、データの精度の変化を計算する。実データパターンデータの差は MAE で計算する。それぞれの家庭の単位時間当たりの平均誤差 $MAE_{All\ Time}$ とピーク時刻の時間のみに着目した誤差 $MAE_{peak\ Time}$ の二つを計算する。

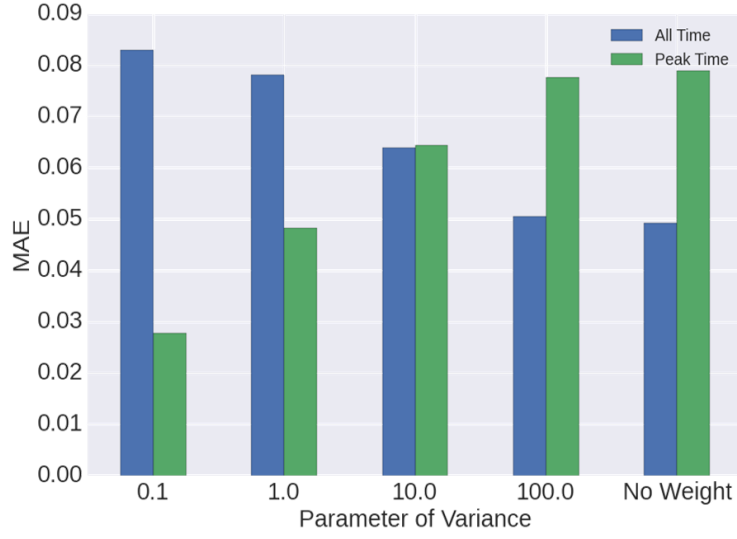


図 5-7 武蔵小杉データセットの MAE

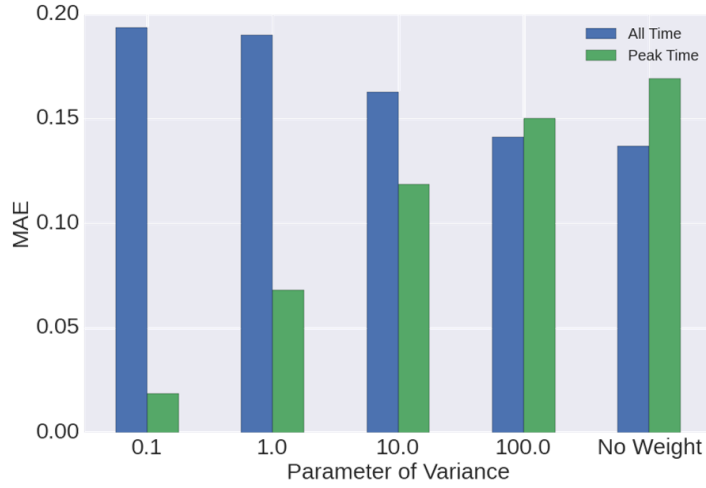


図 5-8 Irish データセットの MAE

$$MAE_{AllTime} = \frac{1}{N_{home}T} \sum_{n=1}^{N_{home}} \sum_{t=1}^T |w_{n,t} - x_{n,t}| \quad (5.3)$$

$$MAE_{peakTime} = \frac{1}{N_{home}} \sum_{n=1}^{N_{home}} |w_{n,p} - x_{n,p}| \quad (5.4)$$

結果を図 5-7, 図 5-8 に示す. ここでは, $\sigma^2 = 0.1, 1.0, 10, 100$ と重みづけなし (No Weight) で全時間の MAE とピーク時刻のみに注目した MAE を計算した結果を表している.

σ^2 が小さい程, ピーク時刻において精度は高く, 一方で全時間の精度は低下した. 特に, Irish データセットの重みづけなしと比べ, $\sigma^2 = 1$ ではピーク時間で 60% 誤差を削減し, 全時間で

第 5 章 評価

40%誤差が増加した. このように重み付けにより全体の精度とピーク時間の精度はトレードオフの関係にある. 全時間で要求される精度を保証したうえで, ピーク時間の精度が最も高い精度となるパラメータ σ^2 を設定する必要がある. 例えば武蔵小杉データセットにおいて, 全時間の誤差精度は 0.1kWh であれば, $\sigma^2 = 0.1$ と設定し, 0.07kWh であれば, $\sigma^2 = 10.0$ と設定すれば全時間平均の要求を満たしつつ, ピーク時刻の精度は最大となる.

第6章 結論

本研究ではスマートメータで送信される電力データを個人のプライバシーを考慮しながらデータ共有する手法を提案した. 自己組織化マップを用いて家庭の使用電力データ共有手法を提案した. 互いに実データを公開することなくパターンデータを生成することが可能となる.

SOM のパターンデータを用いることで, 各企業がそれぞれ独立に匿名化処理を行った場合と比べ, 少ない情報損失度で k -匿名性を実現した. 特に企業数 20 で $k=20$ の場合, 提案手法を用いることで, 独立的に匿名化処理を行った場合と比べ情報損失度約 20%削減を達成した.

また, 提案手法を用いてデータを企業間で共有することによりデマンドレスポンスへの適応が可能となる. 実データと提案手法を用いて生成されたパターンデータの差の評価を行った. 提案手法における設定パラメータ σ^2 が小さい程, ピーク時刻において精度は高く, 一方で全時間の精度は低下した. 特に, Irish データセットの重みづけなしと比べ, $\sigma^2 = 1$ ではピーク時間で 60%誤差を削減し, 全時間で 40%誤差が増加した. このように重み付けにより全体の精度とピーク時間の精度はトレードオフの関係にある. 全時間で要求される精度を保証したうえで, ピーク時間の精度が最も高い精度となるパラメータ σ^2 を設定する必要がある.

参考文献

- [1] 経済産業省, “次世代エネルギーシステムに係わる国際標準化に関する研究会,” 2010. [オンライン]. Available: <http://www.meti.go.jp/report/downloadfiles/g100129d01j.pdf>.
- [2] 電力・ガス事業部, “スマートメーターの最近の動向について,” 第12回スマートメーター制度検討会, 2013.
- [3] 株式会社三菱総合研究所, “海外のスマートメーター及び柔軟料金に関する動向,” 2012年3月12日. [オンライン]. Available: http://www.meti.go.jp/committee/summary/0004668/011_09_00.pdf.
- [4] C. F. E. REGULATORY, “2010 Assessment of Demand Response and Advanced Metering,” 2011.
- [5] 電力中央研究所, “米国における家庭用デマンドレスポンス・プログラムの現状と展望—パイロットプログラムの評価と本格導入における課題,” Y10005, 2011.
- [6] “EFF Advises California PUC on Smart Grid Privacy Protections,” 2010.
- [7] “NIST IR 7628 Vol.2,” [オンライン]. Available: http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol2.pdf.
- [8] 淳. 佐久間, 重. 小林, “プライバシデータ保護マイニング,” 2009-03-01.
- [9] “JAPAN SMART CITY PORTAL,” [オンライン]. Available: <http://jscp.nepc.or.jp/index.shtml>.
- [10] 大京アステージ, “コンソーシアム参加企業との家庭向けサービスの利用を通じた既存マンションにおけるエネルギーマネジメント実証事業を開始,” 2013年8月12日. [オンライン]. Available: <http://www.daikyo.co.jp/dev/files/20130812.pdf>.
- [11] 横浜市温暖化対策統括本部, “横浜スマートシティプロジェクト (YSCP) ビル部門での本格実証～統合 BEMS による大規模ビル間連携実証～,” 2013年3月22日. [オンライン]. Available: <http://www.city.yokohama.lg.jp/ondan/press/h24/130322/130322.pdf>.
- [12] 北九州スマートコミュニティ創造協議会, “北九州市プレゼン資料,” 次世代エネルギー・社会システム協議会 (第14回) - 配付資料, 12 2012 年.
- [13] J. CITY 北九州スマートシティ創造事業, “地域の電力需要に合わせ EV の充電サービス料金も変動,” [オンライン]. Available: <http://jscp.nepc.or.jp/article/jscp/20130129/338446/index.shtml>.
- [14] 川崎市総合企画局スマートシティ戦略室, “「クールシェア」等実証イベントの実施について,” [オンライン]. Available: http://www.city.kawasaki.jp/templates/press/cmsfiles/contents/0000069/69117/150715shoudou_kosugi.pdf.
- [15] 横浜市温暖化対策統括本部, “「横浜スマートシティプロジェクト (YSCP) ビル部門での本格実証～統合 BEMS による国内大規模ビル間連携実証～ 今夏の実施結果 ピークカット 最大 22.8%達成」,” 2013年10月25日. [オンライン]. Available: <http://www.city.yokohama.lg.jp/ondan/press/h25/131024press.pdf>.
- [16] Rakesh Agrawal and Ramakrishnan Srikant, “Privacy Preserving Data Mining,” SIGMOD, pp. 439-450, 2000.
- [17] LSeeney, “k-anonymity: a model for protecting privacy,” International Journal on

- Uncertainty, Fuzziness and Knowledge-based Systems, vol. 10 no. 5, pp.557-570, 2002.
- [18] Ashwin Machanavajjhala, Daniel Kifer and Muthuramakrishnan Venkitasubramaniam, “L-diversity: Privacy beyond k-anonymity,” ACM Transactions on Knowledge Discovery from Data, vol. 1, no. 1, 2007.
- [19] P. Paillier, “Public-key cryptosystems based on composite degree residuosity classes,” Advances in cryptology—EUROCRYPT’ , pp.223-238, 1999.
- [20] T. Kohonen, “The self-organizing map,” Proc. IEEE, vol. 78,no.9, , 1990.
- [21] 江崎 浩. 落合秀也, “インプレス標準教科書シリーズ スマートグリッド対応 IEEE1888 プロトコル教科書,” インプレスジャパン, 2012.
- [22] Data from the Commission for Energy Regulation [オンライン]. Available: <http://www.ucd.ie/issda/data/commissionforenergyregulationcer>.

謝辞

研究を進めるにあたり、日ごろから厚いご指導、ご鞭撻を賜りました西宏章教授にこの場を借りて心より御礼申し上げます。また、小杉駅周辺地区スマートコミュニティ実証事業において、ご協力を頂きました、川崎市スマートシティ戦略室、株式会社大和総研ビジネス・イノベーション、株式会社日本総合研究所のプロジェクトメンバーの皆様に深く感謝いたします。

研究に際し、お世話になりました戸越敏郎様、林泰正先輩、松井加奈絵先生、Erwin Harahap先輩、八巻隼人先輩、Janaka Wijekoon先輩、Rajitha Lakzmal Tennekoon先輩、今西智哉先輩、高際兼一先輩、岡野博典氏、河野翔貴氏、神戸祥夫氏、中村優一氏、福田美桜氏、ハーバス安利氏、野口智道氏、西村公佑氏、池田伸太郎氏、江村祐太氏、塩原寿周氏、子安光樹氏、澤口聡太氏、白木勇矢氏、細江貴博氏、増川健人氏、松井忠宗氏に深く感謝いたします。また、日頃から深い愛情で著者を支えてくださった家族に心より感謝いたします。

本研究は、公益財団法人セコム科学技術振興財団研究助成、文部科学省科学技術研究費補助金基盤研究 B「機能維持性を高める建物・複数機器の協調制御」(24360230)ならびに「コンテナベース・スマートコミュニティインフラの構築と展開」(25280033)、の一環としてなされた。