

企業統計業務の電子投票による効率化提案

Tatsuya HAYASHI,

Graduate School of Media Design, Keio University
Hiyoshi Kohoku-ku Yokohama-city Kanagawa, Japan 223-8526

hayashi@lepidum.co.jp

Ichiro Mizukoshi,

Institute of Information Security, Japan

2-14-1 Tsuruya-cho, Kanagawa-ku, Yokohama-city, Kanagawa, 221-0835, Japan

mizukoshiichiro@gmail.com

ABSTRACT

政策立案などの意思決定には企業活動に関する統計情報(以後企業統計と略す)の活用が欠かせない。例えば、観光政策においては宿泊施設の稼働率、通信政策においてはIPトラフィックの総量推移などが挙げられる。意思決定の精緻化・迅速化は常に望まれており、それに応えるためには企業統計の精緻かつ迅速な提供が必須となる。しかし、現状ではこの作成は手作業に頼る部分が多く迅速な提供を妨げている。この課題らに対し本提案では、1)電子投票の仕組みを統計業務に適応することで効率化を図ること、2)システムコスト削減のために常設の電子投票システム共通基盤を用意することの二点を提案する。

KEYWORD

企業 統計 電子投票

1. はじめに

意思決定においては、その基礎となる情報が必要である。政策分野においては企業活動の状況を把握することが重要であり、企業自体も業界全体の動向を把握することによってより適切な意思決定が可能になる。しかし、収集された情報の有益性には理解を示しつつも、自社の情報を競合他社に知られることに拒否感を示す企業は多い。このため、収集・公開される情報は、市場データのような公開情報か、何

らかの匿名措置を講じて各企業から収集したデータに合計などの統計処理を施したものが公開されていることになる。

筆者らの経験ではこのような匿名化は作業員の守秘能力を前提とした手作業による部分が多い。また統計化においても、該当項目の特異性などを認識し、どう該当情報を丸めるかは作業員の経験に依拠するケースが多数存在する。そこで、本研究では、匿名で企業から収集したデータを統計処理に基づいて公開する作業と電子投票の類似性について検討したのち、常設システムを構築することによるシステムコスト削減について提案する。

2. 先行研究

一藤らは観光政策の基礎データである宿泊施設の稼働状況を把握するために、インターネット上の予約サイトのデータを収集することで、各宿泊施設の稼働率を推定する手法を提案している [1]。この手法は回答者に負担をかけないという点で優れているが、推定であること、予約システムの主たる目的とは異なった利用形態であることに課題がある。

国内ではインターネットトラフィックの集計が2004年から継続的に行われている [2]。この作業は総務省データ通信課を事務局とし、学界の研究者と国内ISPといった、産官学の連携によってトラフィック情報の秘匿性を維持しつつ(個社のトラフィック量が明らかにならないよう)、協力ISP全社の合

計値としてトラフィック量を公開している [3]。筆者らのヒアリングによれば、秘匿性は研究者の独立性に強く依存しているとのことであり、将来的にはシステム化が望ましいと考えられる。

Gritzalis は "Voting systems design criteria" として Authentication, Uniqueness, Accuracy, Integrity, Verifiability, Auditability, Reliability, Secrecy, Non-coercibility, Flexibility, Convenience, Certifiability, Transparency, Cost-effectiveness の 14 項目を挙げており [4]、山崎は日本語で次のように表現している [5]。

1. 投票資格の検査と本人確認
2. 投票の一意性(二重投票の防止)
3. 投票内容の正確性
4. 投票内容の改変不能性
5. 投票内容の検証可能性と追跡可能性^a
6. システム自体の信頼性
7. 個人の投票内容と方法の秘匿性^b
8. 柔軟性(複数の投票手段が選択できる)
9. 利便性(情報リテラシーが不要)
10. 試験可能性(仕様を投票者が確認できる)
11. 透明性(投票に関する全工程の可視化)
12. システムの運用コストが合理的

これらの要件は企業統計において期待される要件と同等であることを後述する。

田上らは投票者の匿名性を保ちつつ、公開の場において集計の正当性が証明可能な技術として無記名式電子投票方式(Electronic Voting Scheme)の利点を認めつつも、投票者・収集者ともに処理が複雑なことを難点としている。これに対して全数調査ではなく標本調査であるインターネット調査においては厳密な集計結果は保証しないものの、簡易な技術で匿名性を保証することの方が重要と主張している。その具体的な手法として Item Count 法 [6]と RR 法(Randomized response techniques) [7]を比較した上で、確率的变化に基づいた RR 法の発展方式を提案している [8]。今回のテーマである企業統計においては、全数調査であっても比較的少数の標本数であることが想定されることから、Item Count 法、RR 法よりは無記名電子投票方式が適切と考えられる。

電子投票の基本的な仕組みに関わる処理の複雑さは如何ともしがたいが、投票機構を常設し投票の機会を増やすことで全体コストの削減が期待できることを本論文で示す。

税所らは、株主総会における議決権行使を例として 1 人が複数票を投票可能な投票 (重み付き投票) の電子化方法の提案とその安全性を評価している [9]。その中で、重みの考慮として以下の 3 要件を上げている (C1-C7 は通常の電子投票に求められる要件)。

C8 (高効率性)重み付き投票を効率良く実現できる。

C9 (不統一行使防止)投票者の投票内容はすべて同一である。

C10 (重み情報の秘匿)重みを秘密情報と見なす場合、投票者の重みに関して、投票結果のみから得られる情報以上の情報を得ることができない。

企業情報(生産量や宿泊者数)を重み情報として捉えた場合、これらの要件を満たすことが重要となる。一般財団法人日本統計協会はサイト上で "民間の統計 業界団体、民間調査機関等の作成する統計" として 200 以上の統計を紹介している [10]。民間統計が網羅されているわけではないが、統計需要を示す一つの証左と言える。

政府統計については独立行政法人統計センターが 2003 年に設置され、総務省統計局所管の国勢調査や消費者物価指数など、わが国の基本となる統計の作成 (製表)、各府省や地方公共団体の委託を受けて各種の統計作成を行い、これらの機関の統計整備を支援している [11]。

^a Verifiability と Auditability を一つにまとめている。

^b Secrecy と Non-coercibility を一つにまとめている。

3. 必要要件

3.1 検討対象

企業統計で扱うデータには様々な種類がある。

表 1 企業統計データの例

データ	データの種類
生産量	量的、連続、間隔尺度
生産台数	量的、離散、間隔尺度
景況感	質的、順序尺度

収集された情報は、個々の情報が(記名・匿名)開示されることもあれば、合算などの統計処理された結果のみが公開される場合もある。統計処理された場合においても、参加者情報(参加数、個別名称など)の公開はケースバイケースである。

本研究では、量的データについては、合算、質的データについてはヒストグラムといった統計処理された結果を公開する場合について検討する。また、参加者情報の公開についての検討は将来の課題とする。

3.2 (電子)投票との比較

山崎があげた 12 要件を CIA で分析すると、

Confidentiality 投票の信頼性 1-6,10,11

Integrity 投票の秘匿性 7

Availability 投票の容易性 8,9,12

と考えることができる。今回の企業統計業務について望まれる要件が、この電子投票の要件とほぼ合致していることは自明であろう。ただし、多くの(電子)投票においては、あらかじめ定められた候補から有権者がどれか一つを選択し、その得票数の合計が公開される。これに対し、本研究の対象では、合計・ヒストグラムといった異なった処理が行われることに留意する必要がある。

3.3 ヒストグラム

質的データは合計などの処理を行うことができないため、その結果をヒストグラムで表現することになる。列挙された質的データを投票における各候補と見なせば、(電子)投票における各候補の得票数を公開するとこと質的データのヒストグラムの公開は

同値となる。このため、ヒストグラムの公開においては電子投票システムの変更は不要である。

3.4 合計

多くの投票システムで有権者は一人一票という公平さを前提としている [9]。企業情報が離散量であれば、各有権者に十分な大きさの投票用紙を与えた信任投票(各自の投票数=各自の離散量)を行うことで合計は可能である。連続量であった場合には、有効桁数を揃えることで、離散量として扱うことが可能になる。

複数回の投票といった非効率性については、税所らが提案している重み付き投票とその 3 要件(高効率性、不統一行使防止、重み情報の秘匿)に基づいた分析が有望と考えられる [9]。しかし、彼らのモデルは、重みが事前に与えられ、不統一行使防止要件によって投票数の自由度が失われているため、企業自身が投票数をコントロールできない。このため、さらなる検討が必要である。

4. システムと運用

4.1 システム化による利点

各府省や地方公共団体であれば、統計センターが委託を行うことが可能である。この統計センターにおいて求められているモデルは、統計を取りたい人と集計者を分離することで実現されている。

これを踏まえて、民間を含め汎用な統計のシステム構築することによって、様々な効率化が期待される。

前掲の通り、統計に求められる要件は、電子投票のシステムにおいて求められる要件と合致しており、汎用的な統計システムへの発展が期待できる。同時に未だ普及していない電子投票において応用事例としてこのような統計システムによって経験値や知見を溜めることは重要だろう。

更にこのシステムが常設されることで、システムが洗練されること、現状の統計処理における属人性の解消、信頼の蓄積などが実現可能である。

4.2 システム化によるリスク

システム化によるリスクは、概ね二つに分類される。ひとつは電子化することによるリスク、もうひとつは分離することによるリスクである。

電子化におけるリスクは、一般の情報システムにおけるリスクと電子投票そのものに関する(現状普及していない要因である)リスクとほぼ同等である。統計処理を分離することによるリスクに関しては、委託処理における委託先の評価や信頼性などがある。これらは、現状の制度においては契約行為などの法的なリスクに還元可能である。

その他、実際に運用コストが現状よりも低減出来るかについてはより詳細な検討が必要だろう。

4.3 システムに求められる運用者の要件

情報システムにおける要件に関しては、前掲の通り電子投票に求められる要件とほぼ同等であり、本稿で特筆すべきはシステムの運用者に求められる要件である。

信頼できる運用者とは、

- 統計処理の過程で不正を行わないこと
 - 入手した情報の目的外利用を行わないこと
- と考えられる。

その為には、不正による受益者でないことが望ましい。一定の中立性と独立性を有した組織が適切である。

これらを実現するためには、組織の条件や情報システム、技術等だけではなく、契約や体制などの社会的な制度とセットになっているシステム(系)として成立することが望まれる。

5. 考察

5.1 平均

企業統計において、全体生産量などのように企業が提供した数値の合計の他に一社当たりの平均値が使用されることも多い。平均値を求めるためには、投票者数が必要となる。一人一票であれば、合計投票数が必要となる。しかし、本提案で検討

した十分な投票用紙の配布では、投票数と投票者数は一致しない。このため、投票者数を別途集計する、あるいは投票内容に数値を許容し、合計を算出するといった拡張が必要になると考えられる。この点については今後の検討課題としたい。

5.2 投票者情報

企業統計には、対象数が比較的少数、企業規模の差が結果にあたる個社毎の影響に大きく影響する、参加者が少なかった場合の結果の取り扱いといった課題と特徴がある。このため、投票権を持つものと、実際に投票を行ったものに関する情報が重要であり、システムの中にどのように組み込んでいくかを慎重に検討していく必要があると考えられる。

5.3 法人番号

法人番号は、法人からの情報提供にあたって、日本国内で法人組織を特定するユニークな識別子として有効だといえる。但し、法人番号は支店などには付与されていないため、そのような統計の取り扱いには考慮が必要である。

5.4 制度

運用者の要件にもあるように、技術や組織の要件だけではなく、社会的な制度による信頼性の向上が重要である。

情報産業の発展と共に、こういった制度とあわせて社会問題を解決するアプローチは増えており、具体的には以下のようなものが挙げられる。

- トラストフレームワーク
- 裁判外紛争解決手続 (ADR)
- 個人情報保護委員会 (PPC)
- 公益財団法人

これらを参考にした制度設計が望まれる。

6. まとめと今後の課題

6.1 プライバシー・プロファイリングの課題

電子的な統計とプライバシーに関しての一例としては、Internetの標準技術を検討する標準化団体であ

る Internet Engineering Task Force(IETF)でも、” Large-Scale Measurement of Broadband Performance” (LMAP) Working group で計測を行う場合に、Privacy の課題が発生することが議論されており、RFC7594: “A Framework for Large-Scale Measurement of Broadband Performance (LMAP)”において IP アドレスをはじめとした Sensitive Information について触れられている。 [12]

企業統計においては参加者数が少ないため、個社の関する情報価値が高く、例えば投票を行わなかったという行為に価値が生じるので、取り扱いには注意が必要になる。一般の投票においても、電子化されることにより、個人の投票の有無の判別が用意になることについても同様に注意が必要になるだろう。

匿名化に関しては、2013 年の総務省「パーソナルデータに関する検討会 技術検討ワーキンググループ」で、

“いかなる個人情報に対しても、識別非特定情報や非識別非特定情報となるように加工できる汎用的な方法は存在しない。”

という報告が出されており [13]、この一点をみても全自動の投票・統計システムの実現には困難が予測されることがわかる。なお、ここでいう匿名加工は平成 27 年に改正された現行の個人情報保護法における匿名加工”ではなく”、改正のための議論における結論であることに留意が必要である。但し、統計などの限定された条件・目的においては個別の技術的な解法は当然あり得、特定領域における統計はその最たるものだといえるだろう。

参考文献

- [1] 一藤裕，曾根原登，“ソーシャル・ビッグデータ駆動の観光政策決定支援システム,” システム/制御/情報 60.4, 2016.
- [2] 総務省 総合通信基盤局，“我が国のインターネットにおけるトラヒックの集計結果(平成 27 年 11 月分),” 2017.
- [3] 長健二郎，健. 福田，“インターネットトラフィックの現状と動向,” 著: インターネット白書 2013-2014, インプレス, 2014, pp. 137-144.
- [4] D. Gritzalis, “Secure Electronic Voting,” New Trends, New Threats, New Options, 7th Computer Security Incidents Response Teams Workshop, 2002.
- [5] 山崎重一郎，“ブロックチェーンの分散台帳を利用した電子投票による集合知の構成: 対称的な非集中型監査と絶対中立的な非可逆的記録 (特集 社会を変えるブロックチェーン技術),” 情報処理: 情報処理学会誌: IPSJ magazine 57.12, 2016.
- [6] D. Judith, C. A. Rachel, H. L. Michael, P. L. Teresa, V. Wendy, E. M. Trena, “The Item Count Technique as a Method of Indirect Questioning: A Review of Its Development and a Case Study Application,” Measurement errors in survey, 1991.
- [7] S. L. Warner, “Randomized response: A survey technique for eliminating evasive answer bias,” Journal of the American Statistical Association 60.309, 1965.
- [8] 田上敦士, 佐々木力, 長谷川輝之, 阿野茂浩, 富浦洋一, “確率変換に基づくインターネット調査手法の解析,” 電子情報通信学会論文誌 B, 92(4), 2009.
- [9] 税所哲郎, 齊藤泰一, 土井洋, 辻井重男, “重み付き投票の電子化とその安全性に関する考察,” 情報処理学会論文誌 44.8, 2003.
- [10] 一般財団法人 日本統計協会, “民間の統計 業界団体、民間調査機関等の作成する統計,” [オンライン]. Available: <http://www.jstat.or.jp/content/民間の統計/>. [アクセス日: 23 5 2017].
- [11] “独立行政法人 統計センター,” [オンライン]. Available: <http://www.nstac.go.jp/index.html>. [アクセス日: 23 5 2017].

- [12] P. Eardley, A. Morton, M. Bagnulo, T. Burbridge,
P. Aitken, A. Akhter, “A Framework for Large-
Scale Measurement of Broadband Performance,”
IETF, 2015.
- [13] 総務省, “技術検討ワーキンググループ報告
書,” 総務省, 2013.