

携帯電話位置情報の感染症対策への活用手法と論点整理

研究分担者 奥村 貴史
(北見工業大学 工学部 教授)

研究要旨

新型インフルエンザ等の国内症例の発生早期においては、濃厚接触者の特定と健康監視が求められる。しかし、同居人や同僚、同級生等の濃厚接触者特定は容易であるものの、初発例が訪問した大規模商業施設や公共交通機関における接触者を効率的に追跡することが困難であった。そこで、患者の居住地域やの移動情報を一般公開し、接触可能性のある住民に保健所への連絡を呼びかけてきたが、多くの住民はこうした情報に関心を示さない。また、感染症患者の住所や移動経路の詳細情報は、プライバシーに深く関わることから、行政機関として詳細に公表することが出来ない。この問題に対して、本研究分担では、現代社会において高い普及率を示すに至った携帯電話の位置情報を用い、患者と住民双方のプライバシーを守りつつ効果的な感染症対策を実現するための検討に取り組んだ。

今年度は、主として、技術面、法制面、倫理面での検討を進めた。その取り組みの結果、当初予定していた「患者、住民双方のプライバシーに配慮した感染症リスク通知手法」の技術詳細の検討に加えて、既に「公開されている患者発生情報のオープンデータ化を通じた感染症発生情報の効率化」と、「国民の生命が危機に晒される一類感染症への対応における位置情報活用」という 3 つの応用について同時に検討することが望ましいことが明らかとなった。

また、2020年初頭より新型コロナウイルスによるパンデミックが発じ、その対応として、アジア各国における感染症対応における位置情報・接触者情報活用技術について情報収集を行うと共に、新たに登場した接触者追跡アプリを含めた技術の分析と類型化を行った。患者、住民のプライバシーへの配慮と感染症対策を両立させる目的で位置情報を活用する手法は、わが国が世界に先駆けて研究を始めた手法である。現在注目されている接触者追跡手法と共に、今後のパンデミック対策の一端を担っていくことが期待される。

今後の実用化には、技術面、法制面、倫理面にいくつかの課題が存在する。今後、携帯電話の位置情報を保有する携帯キャリア、携帯キャリアが保有する位置情報に関するプライバシー保護を監督する総務省総合通信基盤局電気通信事業部消費者行政第二課、地方自治体における健康危機管理の実務を担う保健所長会、国レベルでの感染症対策を所掌する厚労省健康局結核感染症課との調整を進める必要がある。さらに、この新たな技術の公衆衛生への活用に向けて、実際の活用に向けた法制面の整備と、実証実験の実施、携帯キャリア側における実用化支援を目指す必要がある。

A. 研究目的

結核や麻疹、新型インフルエンザに代表される感染力の強い感染症が発生した場合、地方自治体の公衆衛生当局は、濃厚接触者を特定して健康監視する必要がある。そこで、患者の居住地やの移動情報を一般公開し、接触可能性のある住民に保健所への連絡を呼びかけてきたが、多くの住民はこうした情報に関心を示さない。また、感染症患者の住所や移動経路の詳細情報は、プライバシーに深く関わることから、行政機関として詳細に公表することが出来ない。結果として、行政による情報公開は接触者数が多い大規模商業施設や公共交通機関等の利用情報のみになり、潜在的な接触者に効果的に情報提供することが出来なかった。

本研究分担は、携帯電話位置情報を活用することにより、感染者側と住民側双方のプライバシーを守りつつ効果的な感染リスク通知が実現する手法を着想し、2017年より、この問題に取り組んできた。2017年には携帯キャリアとの最初の接触を行い、2018年には全国保健所長会を含む公衆衛生行政内部での調整に進んだ。その後、2019年3月、厚労省結核感染症課にて協議のうえ、本研究班への追加交付という形で、当該技術の実現に向けた技術面、法制面、倫理面の検討と調整実務を進める運びとなった。

このように、本研究分担では、新型インフル対策という文脈で位置情報の感染症対策応用という問題に取り組み始めたが、2020年1月に入り、新型コロナウイルス感染症によるパンデミックが深刻化し、位置情報の感染症対策が急遽世間の耳目を集める状況が生じた。とりわけ、2月に入り、中国、韓国、台湾、マレーシア等、アジア諸国において公衆衛生当局が情報技術や位置情報を積極的に活用する事例の報道が増えた。その後、3月に入ると、接触者追跡のために、位置情報を用いずに接触者情報のみを効率的に扱う手法の提案

とシンガポールにおける実利用が着目されるに至った。わが国の公衆衛生分野において、感染症対策に位置情報や接触リスク広報を活用する研究を進めてきたグループが他に存在しなかったこともあり、必要な情報収集を行うと共に論点整理を行い、研究活動の一環として情報公開や取材対応に取り組んだ。

B. 研究方法

携帯電話の位置情報を感染症対策へと応用していくためには、技術的な妥当性の検証に加えて、法制面での課題整理に加えて、患者や住民のプライバシーに関する倫理的な課題をクリアとする必要がある。とりわけ、携帯電話の活用という点で、携帯位置情報の利用規制を所掌する総務省との調整が求められる。そこで、技術的な研究を進めつつ、総務省総合通信基盤局電気通信事業部 消費者行政第二課と協議し、携帯位置情報の感染症対策活用に向けた技術詳細の整理を進めた。

そのために、携帯位置情報活用に関するコンサルティングを行っている社団法人社会基盤情報流通推進協議会 AIGID(代表理事 東京大学 空間情報科学研究センター 関本義秀准教授)へとコンタクトし、技術詳細を共同して整理した。

また、現在公開されている患者発生プレスリリースの枠組みにおける効率的な感染リスク計算の実現のために、行政データの「オープンデータ化」に深く関わる技術である「セマンティック・ウェブ」研究者へとコンタクトし、患者位置情報の知識表現を効率化するための規格策定に取り組んだ。

さらに、法制面・倫理面での論点整理を進めるため、国内における個人情報保護法制の検討にに取り組んできた研究グループとコンタクトし、論点整理を進めた。

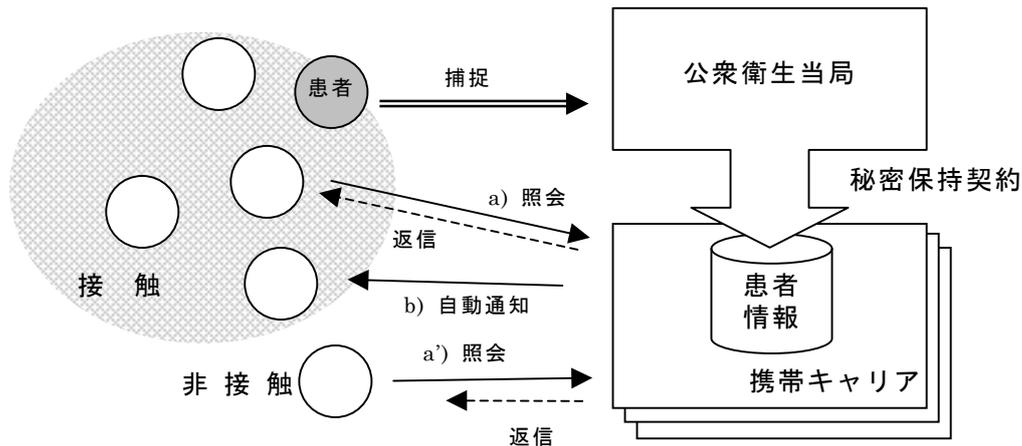


図 1 携帯位置情報を用いた感染症感染リスク通知サービス

C. 研究結果

我々のグループでは、当初、患者と住民の双方のプライバシーを保護しつつ感染症対策を実現する手法(後述する技術分類における「分類 2」)について提案を行い、調整を進めてきた。その後、現在、地方自治体より患者が発生する際にプレスリリースされている公開された患者情報に関して、オープンデータとしての活用を勧める手法(技術分類における「分類 1」)の検討に取り組んだ。また、調整の過程で、国民の生命が危機にさらされる一類感染症において、患者携帯の位置情報を携帯キャリアより行政機関へと提出頂く手法(技術分類における「分類 C」)の検討推進について、公衆衛生行政側より問題提起を頂いた。結果として、位置情報の感染症情報対策応用に関して、3種類の技術について検討した。

まず、「携帯位置情報を用いた接触リスク通知手法」について記す。この手法では、公衆衛生当局は、患者の個人情報の保護に配慮し、携帯キャリアと秘密保持契約を結んだ上で、患者の移動情報を携帯キャリアに提供する(図 1)。携帯キャリアは、顧客からの照会に基づき、当該顧客の端末移動情

報と患者の移動情報を用いて患者との接触リスクを計算し、そのリスクの程度のみを回答する(図 1・a)。感染リスクの通知を事前に希望した顧客に対しては、接触可能性が高い場合、アラートを自動送信する(図 1・b)。この方式により、患者と携帯電話利用者双方のプライバシーを保護しつつ、効果的な注意喚起が実現すると期待された。以上の提案方式の確立に向け、国際会議での発表[Okumura 2019]と計算手法の具体化を目指した技術文書を策定した。

また、提案方式では、公衆衛生当局が積極的疫学調査により得る患者の移動情報を、秘密保持契約(Non Disclosure Agreement: NDA)のもとに携帯キャリアへと提供することになる。患者が発生した際に行われてきた従来のプレスリリースでは、この情報からプライバシーに深く関わる情報を削除したうえで、公共交通機関等他人への接触が生じる移動や滞在のみを公開することが一般的となっている。したがって、従来公開されていない情報を、NDAの下とはいえ行政外部へと提供する形となることから、感染症法や個人情報保護に関する関連法規との整合性の検討が求められていた。一方、既に公開されているプレスリリースのデジタル化は、既に法制的な整理を経て公開さ

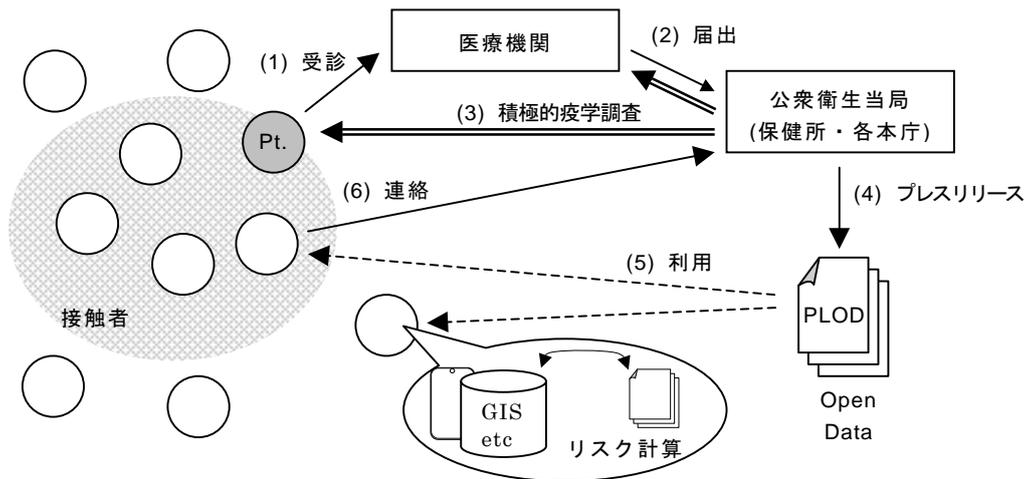


図 2. 患者発生情報のオープンデータ化

れている情報であることから、デジタル化に際した法的なハードルが極めて低い。そこで、オープンデータの要素技術の一つであるセマンティック・ウェブにおける位置情報の知識表現に取り組む研究者と共同し、患者位置情報のオープンデータ化に取り組んだ[Ohmukai 2020]。

また、国民の生命が危機にさらされる一類感染症発生時、患者の協力が得られず適切な入院等の措置が行えず行動の把握が困難となった際、携帯の位置情報を携帯キャリアより行政機関へと提出頂く方法について、法制面の検討を行った。論点として、プライバシー保護、個人情報の保護、通信の秘密の遵守、患者と住民側の人権等に加えて、位置情報の災害対策応用との類似性と差異等、多くの論点が含まれることが明らかとなり、継続して整理を進めている。上述した2件の患者情報公開手法と共に、携帯位置情報を用いた患者情報の収集に関わる論点整理を、資料1として文末に示す。

また、2020年の新型コロナウイルス感染症パンデミック以降、位置情報の感染症対策応用について、アジア諸国を中心にさまざまな手法が報道された。中国においては、患者の購買履歴や公共交通機関の乗車データを元

にして移動経路を把握し、接触者の追跡を行う等の手法が取られた。韓国では、携帯電話の位置情報を用いた患者追跡が行われると共に、移動情報はウェブにて一般公開される。また、自宅待機に際してGPS技術を用いた当局による監視がなされるとされる。シンガポールでは、国土が小さいこともあり、監視カメラ等を駆使した患者と接触者の徹底的な追跡がなされた。研究分担では、パンデミックの初期からこうした各国動向の収集を行い、整理を進めた。その後、米国、ドイツ、シンガポールにおいて、患者との接触可能性のある携帯端末を効率的、高精度に捕捉できる接触者追跡アプリが登場するに至り、有志による国内へのローカライズ作業が進められるに至った。国内でも、蔓延防止のために求められる外出自粛等の過酷さから、アジア諸国に習った位置情報等の技術を用いた対策を求める声がある一方で、プライバシー侵害への懸念から、賛否両論が生じている。しかしながら、そうした議論においては、プライバシー侵害への懸念の低い統計情報の利用にまで懸念の声が出るなど、議論の前提として、2020年に入って移行登場してきたさまざまな技術についての共通理解の基盤が欠

けていることが明らかとなった。そこで、今までの研究の蓄積を元に、位置情報・接触情報を利用した感染症対策技術についての体系的な分類手法を提案した(資料 2)。

D. 考察

公衆衛生分野における携帯位置情報の活用は、研究の歴史が浅い。文献における最初の記載は、2010年におけるハイチ地震の後に、携帯の通話履歴情報を用いて人口移動の統計を作成し、感染症モデルの構築に役立てた事例が最古となる。その後、2015年の韓国における MERS 流行に際して、韓国当局が携帯位置情報を用いた追跡を行ったことが知られていたが、文献上の記載は2017年の報告に言及されているに留まる。その後、2020年に入り、アジア諸国においてさまざまな患者追跡や接触者追跡技術が実戦投入されたが、2020年3月に入り接触者追跡技術に注目が集まるにつれて公開される情報が増していく以前、2020年初頭の段階で研究論文等の形で公刊されているケースはほとんどなかった。

我々の研究の開始時には、個々の携帯端末の位置情報を利用した感染症対策には、韓国のように公権力を用いて患者位置等を捕捉する試みしか存在せず、プライバシー保護上の問題が強く懸念された。一方、我々の提案手法は、そもそもの背景が患者と住民双方のプライバシー保護にあり、携帯キャリア側で行った接触リスク計算の結果が行政には提供されないことにより住民側のプライバシーが保たれる。この方式は、位置情報の感染症対策応用におけるプライバシー保護を進める点で技術的に高い意義があると考えられる。また、携帯位置情報は、GPS を利用したとしても精度に限界があり、接触可能性のある端末を過大に見積もって

しまう。そのため、リスク通知を受け取り保健所側へとコンタクトしてきた住民のなかから、実際に接触リスクの高い者を抽出するための技術を合わせて用意する必要がある。なお、一連の計算より生じる「感染者の潜在的な接触者数統計」は、感染制御に資する貴重な統計データであり、プライバシー情報を含まないことから、携帯キャリア側から行政へと提供可能と考えられる。

なお、2020年3月頃より、位置情報や接触者情報そのものを記録せず、端末間の接触履歴のみを Bluetooth を通じて匿名で収集する各種の手法が登場している。これらの手法は、接触者の検知において我々の提案手法より遥かに高精度であり、また、プライバシー上の懸念が少ない手法として注目される。しかしながら、これらの手法は、我々の手法とは対照的に、利用者数を確保しない限り有効に機能しない欠点がある。また、今後の感染症対策の根幹となりうる技術が海外の「プラットフォーマー」に委ねられることへの危機管理上の課題も存在する。今後の発展に際しては、いずれかの技術が望ましいかという観点ではなく、感染症対策技術における多様性と多層性の確保という観点が重要となるものと考えられる。

なお、一連の研究を通じて、公衆衛生当局による患者情報や接触者情報の収集には課題が多く、今後、この過程がデジタル化されることで効率化されるにせよ、情報形式の標準化が望まれることが明らかとなった。同様に、公衆衛生当局から出る患者情報や接触リスク情報についても、適切な標準化がなされなければ、非効率が生じることが明らかとなった。後者については、今年度の検討を通じてデジタル化に向けた端緒を開いたため、今後、前者を含めた更なる検討が望ましい。

E. 結論

本研究分担では、感染症への対策において携帯電話の位置情報を活用する試みに取り組み、技術面、法制面、倫理面での検討を進めた。その取り組みの結果、当初予定していた「患者、住民双方のプライバシーに配慮した感染症リスク通知手法」の技術詳細の検討に加えて、既に「公開されている患者発生情報のオープンデータ化を通じた感染症発生情報の効率化」と、「国民の生命が危機に晒されうる一類感染症への対応における位置情報活用」という3つの応用について論点整理を行った。

また、2020年初頭より発生した新型コロナウイルスによるパンデミック対応として、アジア各国における感染症対応における位置情報・接触者情報活用技術について情報収集を行うと共に、新たに登場した接触者追跡アプリを含めた技術の分析と類型化を行った。これらの成果については、速報性を重視しネット公開すると共に、新聞社やジャーナリストからの取材対応を通じて社会へと還元している。

感染症危機管理分野においては、これまで、新型インフルエンザ対応としてさまざまな準備を重ねてきた。そのなかでも、患者、住民のプライバシーへの配慮と感染症対策を両立させる目的で位置情報を活用する手法は、わが国が世界に先駆けて研究を始めた手法であり、今後普及するであろう接触者追跡手法と共に、今後の社会における感染制御を担っていくものと考えられる。

その実用化には、技術面、法制面、倫理面にいくつかの課題が存在するため、今後、継続的に検討を進めることにより、新型コロナウイルスへの対応手段としての多様性、多層性の充実に寄与することが望まれる。

F. 研究発表

1. 論文発表

I. Ohmukai, Y. Yamamoto, M. Ito, and T. Okumura, "Tracing patient's PLOD with mobile phones Mitigation of epidemic risks through patients' locational open data", arXiv:2003.06119 (和訳版付)

2. 学会発表

I. Ohmukai, Y. Yamamoto, M. Ito, and T. Okumura, "Tracing patient PLOD by mobile phones: Mitigation of epidemic risks based on patient locational open data", IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE-2020), June 2020.

T. Okumura, "Tracing infectious agents with mobile location information: A simple and effective countermeasure against epidemic risks", 2019 IEEE Global Humanitarian Technology Conference (GHTC 2019), Seattle, October 2019.

3. 資料

資料 1: 携帯電話位置情報の感染症対策活用に際した論点整理

資料 2: 携帯等位置情報の感染症対策応用の現状

資料 3: 感染症対策における携帯電話位置情報の活用手法の技術詳細

携帯電話位置情報の感染症対策活用に際した論点整理

技術的論点

T-1) 位置情報の粒度問題

携帯の位置情報の取得に際して、携帯キャリア側は、基地局在圏情報にはアクセスすることができる。また、この接続基地局情報に基づく経度・緯度情報も利用が可能となっている。しかしながら、基地局情報は粒度が粗く、接触対象者の数が不必要に多くなる懸念がある。この問題の解決策として、基地局電波を用いた測位情報、GPS 情報、WiFi 接続情報を利用しうるケースがある。また、これらの情報と、公共交通機関情報や大規模店舗等の位置情報を用いることで、位置補正を通じた精度向上が図れる可能性がある。この接触リスク計算の精度向上に向けた研究が望まれる。

T-2) 提供情報の形式

患者の位置情報について、行政から携帯キャリアへと情報提供する際、その形式について事前協議を行う必要がある。すなわち、従前のプレスリリースに含まれるような「時間・場所のリスト」形式か、「時間+経度・緯度」形式か。後者である場合、行政における患者の位置情報管理について、ツール等について検討する必要がある。

また、患者の位置情報について、携帯キャリアより行政が情報提供を受ける場合、どのような形式での受け渡しとなり、いかなる実務的課題があるか、検討を要する。

T-3) 顧客カバレッジに関する問題

携帯キャリアには、大手 3 キャリアの他に、大手キャリアの基地局を借用し事業を行う仮想移動体通信事業者(MVNO)があり、その顧客数が全体の 10%程度を占める(2018 年 4 月)。これら MVNO の顧客については、大手 3 キャリアは、契約 ID とその移動情報は得ることができるが、それぞれの契約者のコンタクト先情報は保有していない。したがって、本手法においては MVNO 契約者にリスク通知する手段がない。同様に、海外ローミングにおいても、国内キャリアは、ローミング端末の ID 情報は得られても、海外キャリアが保有する顧客情報が得られず、契約者にコンタクトできない。これら顧客種別に起因する実務的課題について、キャリアに検討を依頼する必要がある。

T-4) 検討・開発コスト負担

本提案の実現に向けて、検討会議の運営と各種検討、サービス開発にコストを要する。その際、携帯キャリア側における検討やサービス開発の費用を公費支出することで、キャリア側の意思決定や対応が迅速となる可能性がある。しかし、行政側は、高額な委託費を負担することは困難かも知れない。この費用負担問題について、何らかの展望を見出す必要がある。

T-5) 保健所側との効率的な接続

濃厚接触者へと送信するアラートには、保健所に相談・連絡するような案内を入れることになる。そこで、保健所側が「詳細な接触状況を確認する仕組み」をオンラインサービスとして整備しすることで、接触可能性のある住民からより接触リスクの高い住民をスクリーニングする必要がある。このサービスを個別自治体が設置することは非効率であるため、国立感染症研究所等の国立機関が設置することが合理的であるが、プライバシーに関わる患者情報を国側が収集しそれを地方自治体に提供する体制やその根拠、予算的基盤について検討する必要がある。

法制面での論点

L-1) 保健所における位置情報収集に関する法的位置づけ

保健所側は、法律に基づいて行う積極的疫学調査の結果として接触者情報を収集している。そうした法的枠組みに基づく情報収集と、今後、「携帯キャリアの情報提供を通じてコンタクトしてくる接触者から収集される情報」や「携帯キャリアから保健所へと提供される患者移動情報」との間には、多少のギャップがある。今後、それぞれの情報収集に関する法的根拠の論点整理を要する。

L-2) 行政機関は、秘密保持契約の下に患者位置情報を携帯キャリアに渡して良いか？

感染者の位置情報は、プライバシーに関わる。したがって、行政が保有する感染者の移動情報を、秘密保護契約の下であれ、携帯キャリアに渡すことが許されるか？という論点がある。なお、感染症法第3章「感染症に関する情報の収集及び公表」には、第16条において、収集した「当該感染症の予防及び治療に必要な情報を新聞、放送、インターネットその他適切な方法により」、個人情報保護に留意しつつ、「積極的に公表しなければならない」としている。これはかなり踏み込んだ表現であることから、秘密保持契約を結んだうえでのキャリアへの情報提供は感染症法の立法趣旨には合致すると考えられる。この解釈について、議論を要する。

L-3) 保健所の有する患者情報の受け渡しに関する手順

患者の移動情報を、誰(公衆衛生行政におけるどの組織)が、誰(キャリアのなかのどの部署)に、どのような形態で渡すかという点について、検討する必要がある。一般的な感染症においては、県設置型の保健所の場合には県庁、政令市の保健所であれば市の担当部局が、情報公開に関する意思決定と実務を担う。一方、今回のような新たな技術を用いた広報については、市町村境界・県境界をまたがった対応が望ましい点においても、県庁や市での対応ではなく、厚労省に情報集約したうえでの意思決定と携帯キャリアへの情報提供が適当となる可能性がある。携帯キャリア側は全国組織であるため、キャリアとしても、各県庁なりからのコンタクトよりも、窓口が厚労省に集約されているほうが、実務上の合理性がある可能性が高い。

L-4) 感染症対策における携帯位置情報の活用に対して、いかに国民の理解を得るか？

感染症対策における携帯位置情報の活用については、「1類感染症対応における止むを得ないプライバシーの侵害」だけでなく、プライバシー侵害のリスクが低い「接触可能性通知サービス」まで、複数の可能性がある。感染症対策において個人の人權侵害が続いた歴史からも、保健所側は、たとえ効率的であっても国民の理解を得ていないサービスを利用することは困難と考えられる。また、「感染症対策における携帯位置情報活用」という表現を目にすると、たとえプライバシーを侵害しない接触可能性通知サービスであったとしても、国民側に「保健所に携帯の移動情報が渡る」との誤解を受ける可能性がある。そこで、携帯位置情報を感染症対策に応用するのに際して、いかにして国民理解を得るかという論点が生じる。

倫理に関わる論点

E-1) 住民への接触可能性通知サービスは、倫理的に許容されるか？

接触リスク通知サービスは、感染者、住民双方のプライバシーに配慮した手法であり、現行のプレスリリース手法以上に不特定多数の接触者に注意喚起を行うことができる。これは、住民側にとって利益となる手法であり、個人にとって不利益となる犯罪捜査における位置情報利用等とは条件が大きく異なる。今後の活用の際には、関係した倫理的な論点を整理したうえで、国民の理解を得ていく必要がある。

E-2) 携帯位置情報を利用した濃厚接触者リスト・統計の作成は、許容されるか？

携帯キャリア等位置情報を保有する事業者は、行政が提供する患者位置情報を利用することで、患者との接触の可能性の高い端末リストや統計を作成することが出来る。これらデータの作成と行政への提供が倫理的に許容されるか、検討を要する。

E-3) 接触可能性通知サービスは、一類感染症以外でも利用が許容されるか？

エボラ出血熱等において接触可能性通知サービスを稼動することは、関係組織の合意が得られる可能性が高い。また、接触可能性通知サービスは携帯利用者のプライバシーを侵さないため、新型コロナウイルス感染症等、社会的インパクトの大きい疾患においても稼動が望まれる。一方で、国民理解が得られないままの運用は、誤解や混乱を生じる可能性があることから、一類感染症以外のいかなる状況であれば稼動が許容されるか、検討を要する。

E-4) 行政は、患者の携帯番号を元に携帯キャリアより位置情報の提供を受け取れるか？

公衆衛生行政側には、一類感染症への対策として、患者位置の効率的な把握を可能とする情報収集手段の用意が望ましいという見解がある。携帯キャリア側も、犯罪捜査における令状ベースの情報提供の枠組みを参考に体制整備することは可能との見解がある。国民の生命が危険に晒される一類感染症における効率的な対応体制の整備に向けて法的、実務的な整理が望まれる。

携帯等位置情報の感染症対策応用の現状

分類	プライバシー侵害リスク	利用手法	利用位置情報(患者・接触者側)				プライバシー情報(行政に対して)			学術文献初記載	実戦投入時期
			携帯電話在圏情報	WiFi基地局接続情報	CPS情報	その他	非公開	限定公開	公開		
0	低	匿名化された携帯位置情報の統計利用	○	△	○	-	○	-	○	要調査	日本(2020.4)、他
A		聞き取りによる移動情報収集	-	-	-	-	-	△	-	要調査	現在、各国の標準運用
B		(自由意志による)接触者追跡アプリの利用と情報提供	精度不足	△	精度不足	○	-	○	-	未記載	米国(2020.3) 欧州?(2020)・シンガポール(2020) 日本(有志チームが準備中)
C	高	携帯位置情報等を用いた患者・接触者追跡	◎注1	○	○	△	-	△	○	[Kim 2017]	韓国政府(2015・MERS) オーストラリア(2020) 中国(2020) 米国(検討中)
D		自宅隔離対象や移動制限者の携帯位置情報監視	◎注2	-	◎注3	-	-	-	○	未記載	台湾(2020) 韓国(2020)

分類(※)	プライバシー侵害リスク	利用手法	利用位置情報(住民側)				プライバシー情報(住民に対して)			学術文献初記載	実戦投入時期
			携帯電話在圏情報	WiFi基地局接続情報	GPS情報	その他	非公開	限定公開	公開		
0	低	患者発生情報のプレスリリース(非構造化、フリーテキスト)	-	-	-	-	-	-	○	不明	現在、各国の標準運用
1		オープンデータ化された患者発生情報(携帯位置情報を利用した感染リスク計算)	△	△	◎注4	-	-	-	○	[Ohmukai 2020]	日本: 複数チームが準備中
2		携帯位置情報を用いた患者接触リスク通知	○注5	○	○注6	-	-	-	○	[Okumura 2019]	中国: 濃厚接触検出器(2020.2) 日本: 未定
3	高	患者位置情報の詳細開示	-	-	-	-	-	-	-	未記載	韓国(2020) マレーシア(2020)

※ 情報の利用形態を具体的に示すため、右の添え字を付しても良い。(p) 個人にパーソナライズした情報提供、(m) 地図にマップした情報提供、(s) 統計処理結果の情報提供

- 注1 1類感染症を対象に、令状ベースで通信事業者等に位置情報開示を求めるとは、法的根拠がクリアでない
- 注2 わが国においては、実現のハードルが極めて高い
- 注3 国内に於ける旅行者を対象に、位置情報アプリケーションのインストールを義務付ける、ないし、同意の下にインストールさせる手法
- 注4 住民側は自らのGPS位置情報を記録するアプリケーションを用い、公開されている感染者データとの接触を計算する
- 注5 携帯キャリアの保有する在圏情報を利用する点で、キャリアによって端末契約者の利用許諾を得る手順が必要となると共に、目的外利用の懸念が生じる
- 注6 住民側が自らのGPS位置情報を記録するアプリケーションを用いる点で、通信事業者と比して、事業者側にとって実現のハードルが低い

利用例: 日本(2020年4月現在) A0、中国 C2、韓国 D3 黄色: 谷口班検討範囲: 分類1p、分類1m、分類2p、分類C(1類感染症対象)

感染症対策における携帯位置情報の活用に向けたデータ処理概要

令和 2 年 3 月 6 日

文書の目的と概要

結核や麻疹、新型インフルエンザに代表される感染力の強い感染症が発生した場合、地方自治体の公衆衛生当局は、濃厚接触者を特定して健康監視する必要がある。しかし、行政による情報公開は接触者数が多い大規模商業施設や公共交通機関等の利用情報のみに留まり、潜在的な接触者に効果的に情報提供することが出来なかった。この問題に対し、本報告書は、2019年10月に発表された「携帯電話の位置情報を用いて患者と携帯電話利用者双方のプライバシーを保護しつつ効果的な感染症対策を実現する手法」の具体化に取り組んだ。

1章では、本報告書の元となる感染症対策手法を概説した。2章では、感染リスク計算に活用する位置情報について、その種類とその概要、精度等の特性について整理した。3章では、携帯位置情報を用いた接触リスク評価の具体的な処理方法について概説した。4章では、それを具体的に実装し、接触リスク通知サービスとして実現するためのサービス形態や運用方法、関係法令を整理した。5章では、コストの低廉化と計算精度の向上を目指し、接触リスクの計算を第三者機関が実施する将来的な可能性について記した。

目 次

1. 感染症対策における携帯位置情報の活用	4
2. 感染リスク計算に活用しうる位置情報	6
2.1 基地局に係る位置情報	6
2.2 GPS 位置情報	7
2.3 Wi-Fi 位置情報	9
2.4 携帯キャリア等の保有する位置情報の商用利用例	9
2.5 感染リスク計算に活用しうる位置情報の取得における課題	11
3. 携帯位置情報を用いた接触リスク評価	12
3.1 携帯キャリアにおける位置情報の保存と前処理	12
3.2 位置情報に基づく接触リスクの概算	13
3.3 位置情報を利用した接触リスクの高精度推計	14
3.4 接触リスク評価に関する検討事項	16
4. 接触リスクの通知サービス	19
4.1 接触リスク通知のプル型サービス	19
4.2 接触リスク通知のプッシュ型サービス	22
4.3 位置情報活用と関係法令	23
5. 接触リスク通知サービスの外部委託	27

1. 感染症対策における携帯位置情報の活用

結核や麻疹、新型インフルエンザに代表される感染力の強い感染症が発生した場合、地方自治体の公衆衛生当局は、濃厚接触者を特定して健康監視する必要がある。そこで、患者の居住地域や移動情報を一般公開し、接触可能性のある住民に保健所への連絡を呼びかけてきたが、多くの住民はこうした情報に関心を示さない。また、感染症患者の住所や移動経路の詳細情報は、プライバシーに深く関わることから、行政機関として詳細に公表することが出来ない。結果として、行政による情報公開は接触者数が多い大規模商業施設や公共交通機関等の利用情報のみに留まり、潜在的な接触者に効果的に情報提供することが出来なかった。この問題に対し、携帯電話の位置情報を用いて患者と携帯電話利用者双方のプライバシーを保護しつつ効果的な感染症対策を実現する手法(図 1-1)が提案されている [1]。

まず、該当する患者が医療機関を受診した際(1)、感染症法に基づき、医療機関は保健所に対して患者情報の届出を行う(2)。保健所は、届出を受けて、積極的疫学調査として患者関係情報の照会・収集を行う(3)。この収集情報には、患者氏名、住所、既往歴やリスクファクタ等の症例情報に加えて、接触者情報が含まれ、極めて機微な個人情報が含まれる。

公衆衛生当局は、患者の個人情報の保護に関する秘密保持契約を結んだ携帯キャリアに、「患者の移動情報」のみを提供する(4)。携帯事業者の顧客のうち、感染症リスクを知りたい者は、携帯キャリアに照会する(5-a)。携帯キャリアは、顧客からの照会に基づき、当該顧客の端末移動情報と患者の移動情報を用いて患者との接触リスクを計算する(6)。その上で、携帯キャリアは、リスクの程度を顧客に返信する(7-a、7-a')。感染リスクの通知を事前に希望した顧客に対しては、接触可能性が高い等のリスクが生じている際、その旨を通知する(7-b)。通知メッセージには、最寄の保健所情報等を含めることにより、相談を促す。感染リスクを告げられた患者は、自らの意思により保健所への相談を行う(8)。

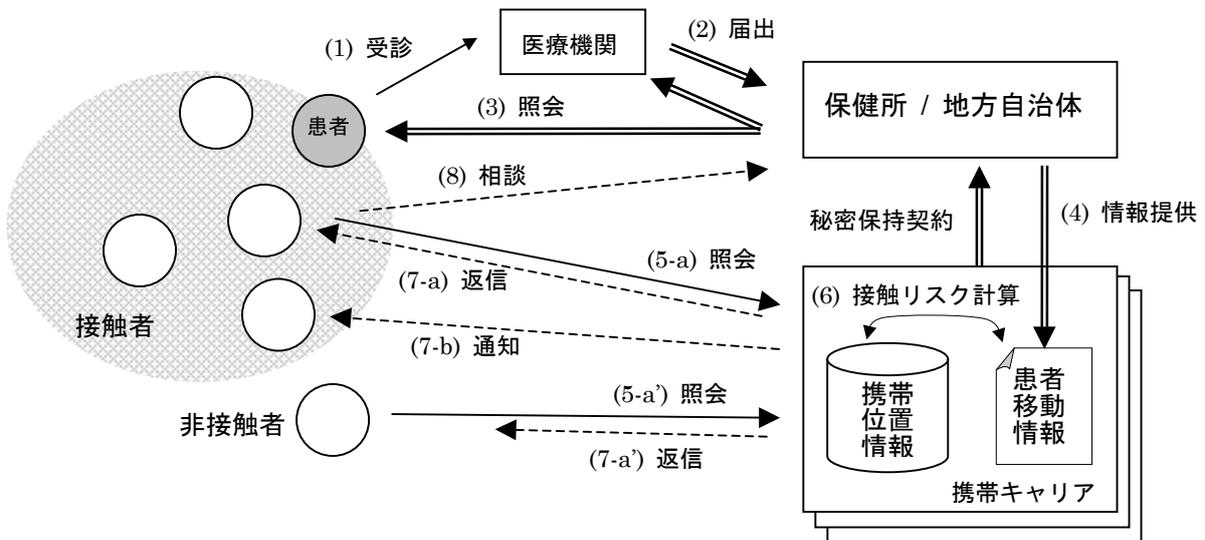


図 1-1. 携帯位置情報を用いた感染症感染リスク通知サービス

以上の手順を踏むことにより、患者情報を公開することなく感染リスクを簡便に周知することが可能となる。また、携帯契約者情報について、携帯キャリアが行政機関に情報提供を行うことを避け、契約者のプライバシーと通信の秘密を守ることが出来る。

以下では、本方式に基づく感染症対策における携帯位置情報の活用に向けて、ステップ(5)～(7)にあたる携帯事業者(以下、携帯キャリア)内部における位置情報処理について、その技術詳細の整理を行う。なお、ステップ(6)については、携帯電話の基地局を用いた位置情報処理の他に、Wi-Fi 基地局情報を用いた処理(6')と、携帯用アプリケーションが保有する GPS 情報(6'')を用いたバリエーションが考えうる。これらの位置情報処理についても、携帯キャリアが扱うケースが考えうるため、本文書にて合わせて取り扱うものとする。ステップ(4)については別文書にて詳述するものとし、本文書においては、当該データの適切な前処理によって任意の時点での患者の移動情報が緯度・経度にて与えられることを前提する。

2. 感染リスク計算に活用しうる位置情報

2.1 基地局に係る位置情報

基地局に係る位置情報(図 2-1)については、文献[2]で次の様な整理がなされている。

基地局に係る位置情報は、携帯電話事業者等の電気通信事業者が通話やメール等の通信を成立させる前提として取得している情報のことであり、位置登録情報と個々の通信の際に利用される基地局の位置情報に分けられる。

位置登録情報とは、移動体端末が着信等を行うために、移動体端末がどの基地局のエリア内に所在するかを明らかにするため、移動体端末がエリアを移動するごとに基地局に送られるほか、あるエリア内でも定期的に基地局に送られる情報をいう。具体的には、基地局の識別番号、端末の識別番号、取得日時等によって構成される。実際に通信している際に用いられる情報ではないため、携帯電話事業者等においては個人情報として扱われる。精度は基地局単位であり、概ね数百メートル単位である。

他方、「個々の通信の際に利用される基地局情報」は、携帯電話事業者等においては通信の秘密として取り扱われ、基地局の識別番号、通信の発信元の識別番号、通信の発信先の識別番号、通信日時によって構成される。精度は位置登録情報と同じく基地局単位である。

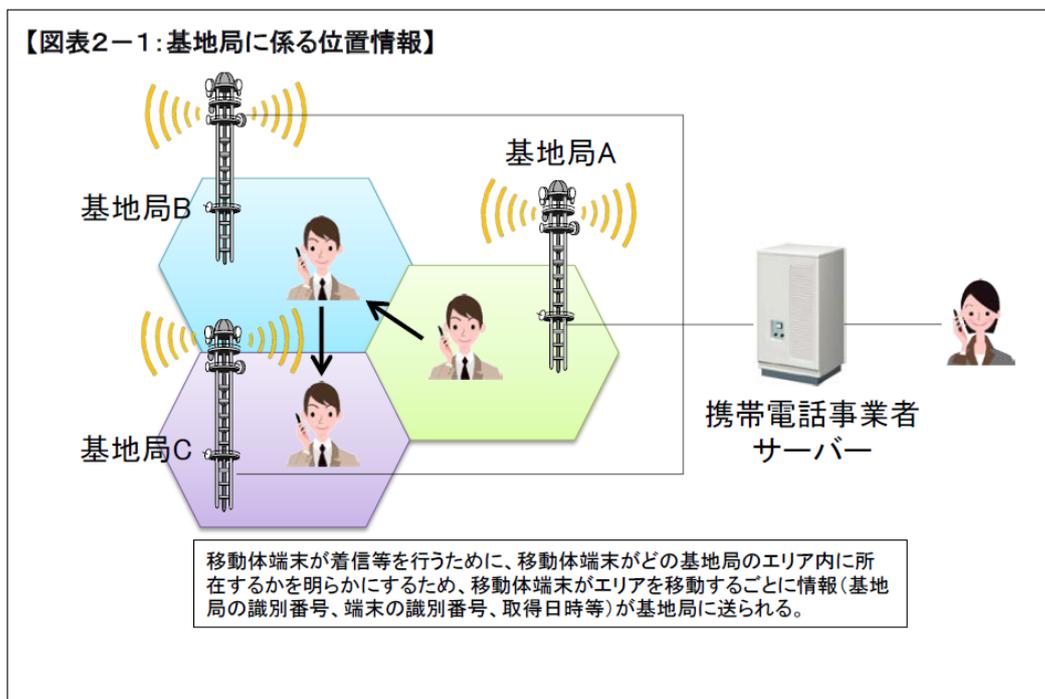


図 2-1. 基地局に係る位置情報 [2]

これらの基地局に係る位置情報を管理するノードは、3G ネットワークでは HLR (Home Location Register)、LTE では HSS (Home Subscriber Server)と呼ばれる。これらは、加入者回線の MSISDN (Mobile Station ISDN Number; 電話番号)、IMSI (International Mobile Subscriber Identity; 加入者を識別するための国際番号)などの情報を保持し管理する回線データベースとしての機能、MME (Mobility Management Entity)などのネットワーク内の各ノードよりの問い合わせに応答し接続や機能の利用可否について加入者回線を認証する機能、音声呼や SMS などのサービスを移動通信ネットワークで実現するため端末が収容されているネットワークや交換機の情報リアルタイムに保持する在圏管理の機能、SIM カードと対になる情報を格納し通信内容の秘匿に必要な暗号化機能を他のノードと併せ提供する機能等を有する[3][4][5]。移動体通信ネットワーク内の HSS とその他のノードが基地局に係る位置情報をやりとりするインターフェースは 3GPP の仕様にまとめられている[6]。

通信方式によらず、携帯キャリアの通信網から得られる位置情報は、接続した携帯電話基地局の位置情報の配列となる。携帯電話基地局のカバー範囲は、人口分布や設定場所の地形などで地域差があり、1km 程度から数十 km 程度までと幅がある。さらに、必ずしも最近隣の基地局に接続しているとは限らない。このため、基地局による方法では、大まかな移動傾向を読み取れる程度であり、この位置情報をもとに必ずしも本人の移動経路を把握できるわけではない。具体的な移動情報を推測するためには、道路や施設などへのマップマッチング等の前処理が必要となる。

なお、LTE では、OTDOA (Observed Time Difference Of Arrival)や E-CID (Enhanced-Cell ID)といった測位方式が導入されており[7]、50m より小さい精度での位置情報が得られるとの報告がある[8]。しかしながら、全ての端末をカバーしているわけでない点で、汎用性に欠けている。一方、基地局の在圏情報を利用した位置情報は、基地局が日本全国の多くの領域をカバーしている網羅性があるため、都市間などの広域での移動の把握に期待ができる。地下鉄駅などの屋内空間においては、個別のアンテナがあることから、閉鎖空間であれば接続したアンテナ ID 等から位置を特定することは比較的容易と考えられる。ただし、駅のどの辺りかまでの具体的な位置情報を得られるほどの空間分解能を得ることは困難である。

なお、携帯キャリアの保有する情報は、日本では「通信の秘密」となる可能性が高いため、利用に際して事前に本人の明確な同意を得ること(オプトイン)が必要である。携帯キャリア内でデータ処理を行う場合であっても、携帯電話サービスでの利用の目的外になるため、本人同意の取得が必要になると考えられる。

2.2 GPS 位置情報

GPS 位置情報(図 2-2)については、文献[2]において、次の様な整理がなされている。

GPS 位置情報とは、複数の GPS 衛星から発信されている電波を携帯電話等の移動体端末が受信して、衛星と移動体端末との距離等から当該移動体端末の詳細な位置を示す位置情報である。緯度経度情報、端末の識別情報、取得日時等で構成される。GPS 位置情報は、個々の通信を成

立させるために必要な情報ではない上に、電気通信事業者が通信を成立させる前提として取得するものでもない。しかし、その精度は緯度経度単位(数メートル～数十メートル単位)であり、基地局に係る位置情報と比べて移動体端末の詳細な位置を示すことが可能である。電気通信事業者においては、GPS 位置情報は個人情報として取り扱われることが多いが、より精度の高い位置情報であることから、基地局に係る位置情報と比べ、高いプライバシー性を有するとされている。

GPS 位置情報の取得には、利用者端末へのアプリのインストールが必要となる。そのため、それを活用できるのは、アプリ利用者に限られる。アプリで利用している位置情報は、GPS 以外にも Wi-Fi や BLE(Bluetooth Low Energy)などもあり、屋外の移動情報だけでなく、屋内の移動情報についても取得できる可能性がある。ただし、屋内については、屋外と比べて精度のばらつきが大きいと考えられる。

位置情報を取得しているアプリ(サービス事業者)は多数あり、取得方法やデータ構造などサービス事業者ごとに異なることに留意が必要となる。このため、本レポートの検討対象となる感染症対策での利用においては、その収集、データ処理に対して、各サービス事業者との連携方法、連携後のデータ処理方法、データの標準化など、検討が必要と考えられる。

移動体端末そのものの位置情報であるため、自宅、勤務先、立ち寄り先などについて推定できる可能性がある。このため、前項でも記載したように、利用に際して事前に本人の明確な同意を得ること(オプトイン)が必要である。

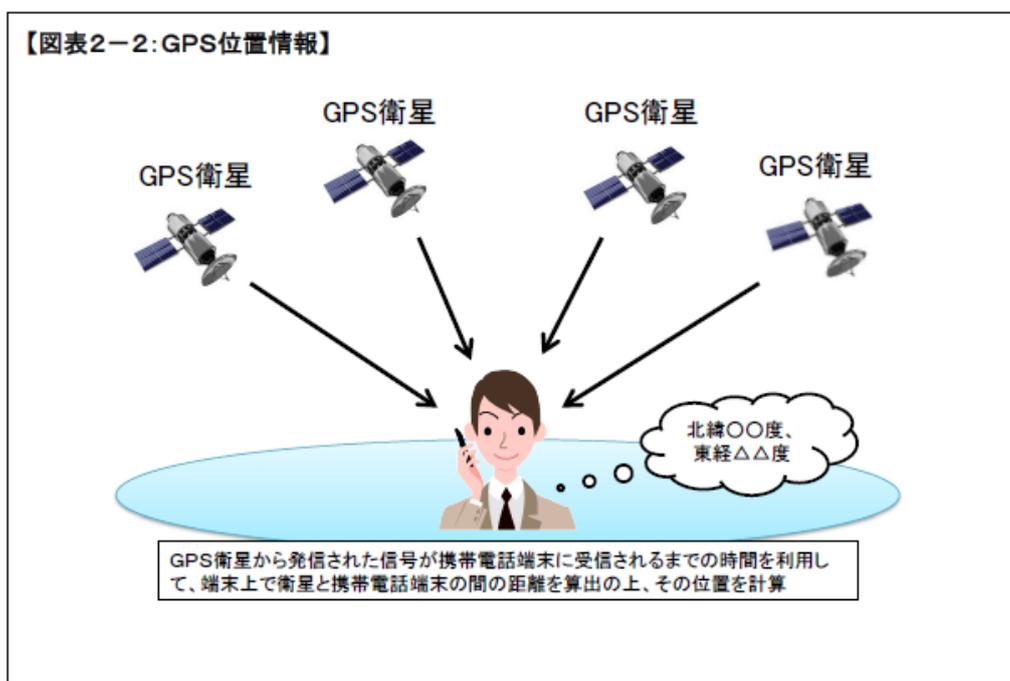


図 2-2. GPS 位置情報 [2]

2.3 Wi-Fi 位置情報

無線 LAN 位置情報の利用を考える場合、次の 2 通りが想定される。このうち 2) は、前節の携帯端末アプリで取得された GPS 情報とデータ管理上等価とみなしうるため、ここでは 1) について述べる。

- 1) NTT-BP (<http://www.ntt-bp.net/>) や Wi2 (<https://wi2.co.jp/jp/>) など、バックエンドの Wi-Fi インフラを提供している事業者の記録しているアクセスログを利用するケース
- 2) 整備された無線 LAN 位置情報のデータベースを利用する、いわゆる Wi-Fi 測位 (Android, iOS の SDK に含まれる WiFi 測位機能等) を利用するケース

Wi-Fi 位置情報については、文献[2]において、次の様な整理がなされている。

Wi-Fi 位置情報とは、Wi-Fi のアクセスポイントと移動体端末間の通信を位置情報の測位に応用することによって、利用者によるインターネット接続の前後を問わず取得される位置情報である。例えば、プローブクエスト(端末が、周囲にある接続可能なアクセスポイントを探すために発信する信号。この信号の中にはMACアドレスが含まれる。)の強度・時間等の情報を用いることで、アクセスポイントのエリア内における端末の相対的位置を推定するものがある。屋内での測位が難しい GPS 位置情報と異なり、屋内外を問わず利用することが可能な Wi-Fi 位置情報は、精度がアクセスポイント単位(数メートル~数十メートル単位)と高いこともあり、大型店舗等で活用される事例が見受けられる。取得される情報としては、アクセスポイントのエリア内の相対的な端末の位置、端末の MAC アドレス、取得日時等がある。

大まかな位置取得の仕組みは、携帯電話基地局と同様である。無線 LAN アクセスポイント (Wi-Fi-AP) に接続、或いはスキャンしたタイミングで、利用者の端末 ID と合わせて Wi-Fi-AP の設置場所の位置情報が記録される。このため、空間分解能は Wi-Fi-AP の分布に依存する。Wi-Fi-AP がないところのデータ取れないため、空間的な網羅性としては、市街地などの Wi-Fi-AP の設置が進んでいるエリアに限定される。Wi-Fi 信号の到達距離を考慮すると、携帯電話基地局に比べれば粒度は細かいと考えられる。同一の閉鎖空間であれば、同じ場所にいることをアクセスした Wi-Fi-AP の ID から判断することも可能と考えられる。

2.4 携帯キャリア等の保有する位置情報の商用利用例

携帯電話ネットワークは電話やメールなどをいつでもどこでも利用できるように、各基地局のエリアごとに所在する携帯電話を周期的に把握している。例えば NTT ドコモはこの仕組みを利用して携

携帯電話の台数を集計し、地域ごとにドコモの普及率を加味して人口を推計したデータをモバイル空間統計として販売している[9][10][11]。モバイル空間統計において、NTTドコモは携帯電話利用者の個人情報に対し3段階の処理(非識別化処理、集計処理、秘匿処理)を行い、プライバシーを慎重に保護している(図2-3)。

このような携帯キャリア等の保有する位置情報の商用利用事例を下記に示す。NTTドコモのモバイル空間統計以外は、アプリによる位置情報(GPS+Wi-Fiなど)を採用している事例となっている。いずれも、個人が特定できない形式に加工されており、位置情報取得の際には、位置情報等のデータの利用、ならびに位置情報等データを個人が特定できない形式に加工した上で第三者に提供することについて、利用者より同意が得られたものとなっている。

- ドコモ:モバイル空間統計(携帯電話基地局在圏情報)
- ゼンリンデータコム:混雑統計(アプリ取得のGPS位置情報)
- KDDI:Location Analyzer(アプリ取得のGPS位置情報)
- Agoop:(アプリ取得のGPS位置情報)
- ブログウォッチャー:profile passport(アプリ取得のGPS位置情報)
- Wi2:(WiFiアクセスログデータ)

前処理の内容については、各社で説明に差があるものの、概ね類似の処理がなされている。[9]、[12]をもとに、以下に整理する。

①非識別化処理:

- 電話番号などの個人を識別できる情報は不使用、また生年月日を年齢層に変換するなどの情報の要約を実施(NTTドコモモバイル空間統計)。
- 生活圏(自宅および職場等の日常行動範囲)の移動履歴を削除(Wi2)

②集計処理:

- 性別・年代別などの属性別に携帯電話の台数を数え、さらに、シェア率などから他社の状況についても推計して加味(NTTドコモモバイル空間統計)
- 位置情報を地域メッシュに加工して統計データとして集計(Wi2)

③秘匿処理:

- 少人数エリアの数値を除去(NTTドコモモバイル空間統計)
- データに紐づくID情報をハッシュ化(生活圏を出てから戻るまでごとにIDを変更、同一IDが付与されるのを最長で1日に制限)(Wi2)

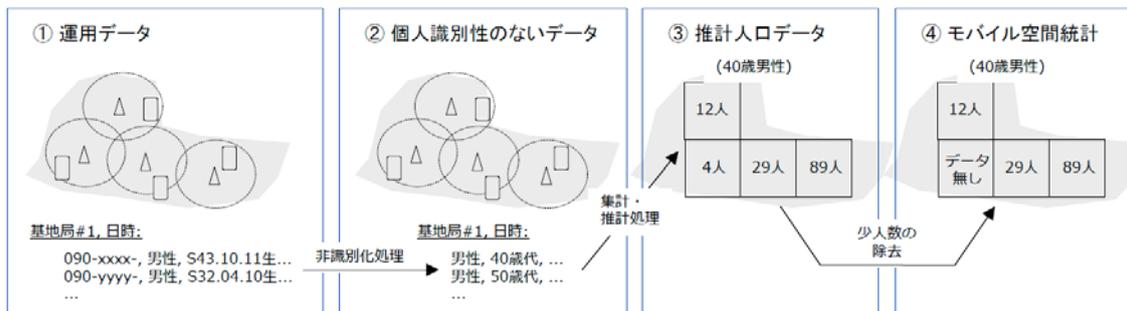


図 2-3. モバイル空間統計の 3 段階処理^[11]

2.5 感染リスク計算に活用しうる位置情報の取得における課題

感染リスク計算に活用しうる携帯端末の位置情報の諸種については、以上のように整理することができる。表 2-1 に、それぞれの地域網羅性、位置精度、サンプリング間隔について整理する。

こうした位置情報を利用する場合には、実務上、いくつかの課題が存在する。まず、携帯キャリアの保有する位置情報は、日本においては「通信の秘密」となる可能性が高いため、利用に際して本人の明確な同意を得ることが必要となる。携帯キャリア内でデータ処理を行う場合であっても、携帯電話サービスでの利用の目的外になるため、本人同意の取得が必要になると考えられる。この同意取得については、第 4 章にて改めて議論する。

また、各キャリアは自社のユーザー情報は把握できるものの、MVNO 端末については、位置を把握することはできても契約情報を有しておらず、連絡に用いることができる情報を保有していない点が挙げられる。同様に、海外旅行者については、ローミングを利用する場合と日本国内においてはプリペイド SIM を利用するケースがありうる。これらについても、端末 ID を用いた位置情報の取得は可能と考えられるが、各契約情報を有していないため、連絡に用いることができる情報を保有しない。

登録なしに接続可能な公衆無線 LAN については、接続した個人を特定することが困難であるため、継続的な追跡に用いることが困難と考えられる。

表 2-1. 感染リスク計算に活用しうる位置情報

	地域網羅性	位置精度(空間分解能)	サンプリング間隔(時間分解能)
基地局に係る位置情報	日本全国	低い(数 km～数十km)	基地局への通信毎 基地局の移動毎
GPS 位置情報	屋外	高精度(数 cm～数十m)	アプリ依存
WiFi 位置情報	都市部の屋内外	中程度(数 m～数百m)	アプリ依存

3. 携帯位置情報を用いた接触リスク評価

携帯位置情報を用い感染リスクを評価していくためには、感染者の位置情報と携帯キャリアが保有する契約端末の位置情報とを用いた接触可能性の計算を行う必要がある。接触リスク評価に際しては、リスク評価を行う端末数、患者数に加えて、計算対象とする期間の長さや時間分解能、リスク評価に用いる地理的な分解能に比例して計算量が増加すると考えられる。この計算量の削減のため、i) 位置情報の保存と前処理、ii) 接触可能性の低い端末の除外処理、iii) 接触可能性のある端末の抽出処理、という3段階の計算を提案する。本章では、この3段階の接触リスク計算について概説する。

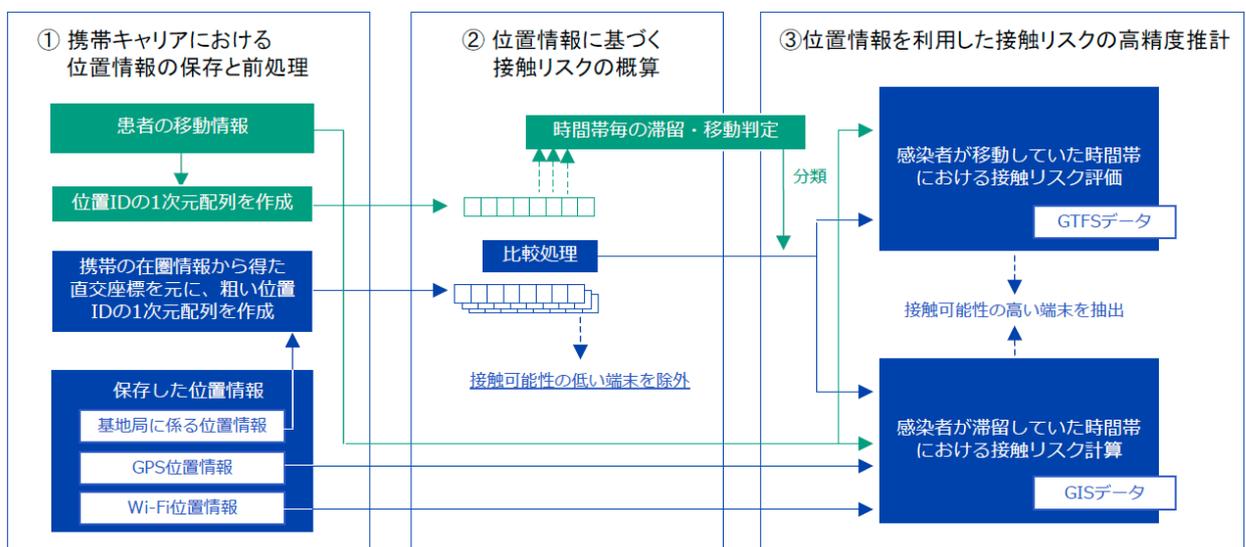


図 3-1. 携帯位置情報を用いた接触リスク評価の手順

3.1 携帯キャリアにおける位置情報の保存と前処理

携帯位置情報を用いた接触リスク評価を行うためには、携帯キャリアを始め各種位置情報保有者は、感染リスクのある一定期間、端末の位置情報を保持しておく必要がある。この位置情報としては、最も粒度の粗い携帯電話の基地局上の在圏情報を基本としつつ、接触リスクの詳細計算のため、Wi-Fi位置情報等のより詳細なデータも保存することが望ましい。この位置情報の保存の際、モバイル空間のような非識別化処理を行うと、端末毎の接触リスク評価を行うことができない。そのため、保存データには端末の識別番号を含め保存する必要があることから、プライバシー上の配慮を要することになる。

保存期間としては、医学的な観点からは、「主要な感染症において患者が発症し病院にて診断がなされ、公衆衛生当局より注意喚起がなされるまでの時間差」を考慮し、2週間前後は保存され

ていることが望ましい。ただし、保存期間が長いことは、プライバシー保護の観点から問題があることに加えて、あまりに長期間の保存を行うためにはストレージへの投資と運用コストが高額となる懸念がある。そのため、記憶容量としては必要に応じて 1 ヶ月程度の記憶容量を確保したうえで、平時は 1 週間程度の保存を行い、公衆衛生当局がアナウンスする感染症の警戒レベルに応じて、2～4 週間に保存期間を延長する措置を取る手が考えられる。

接触リスクの評価を行う際には、まず、携帯の在圏情報を元に、緯度経度へと変換したデータを生成する。その上で、緯度経度の 2 次元情報から仮想的なセル ID に次元を圧縮する。その際、メッシュの粒度は、予備実験を通じて決定するものとする。また、時間軸として、別途定める時間分解能を採用し、各端末の位置情報を 1 次元配列へと圧縮する。このデータ構造を用いることにより、「接触リスク概算」(3.2 章)は、患者側の移動情報を保存した 1 次元配列との比較処理へと単純化されることになる。

3.2 位置情報に基づく接触リスクの概算

携帯キャリアが保有する膨大な契約者の端末移動情報のなかから、患者と接触した可能性のある携帯端末をリストアップするためには、膨大な計算量を要する可能性がある。この問題は全契約端末を対象とした計算を行わないことにより緩和されるが、将来的に契約端末の相当数を対象とした接触リスク計算を行う必要が生じた際には現実的に克服すべき課題となる。そのため、制度の設計段階から計算量の適切な削減を考慮した手法を用意する必要がある。

計算量を削減するためには、まずは低い分解能において接触可能性の低い端末を除外することで、詳細な接触可能性の計算を接触可能性の高い端末に限定することが望ましい。そこで、3.1 章で用意した各端末の位置情報(1 次元配列)と、患者側の移動情報より作成した位置情報との比較処理を行う。ここで用いる患者側の移動情報は、図 1-1 の(4)にて公衆衛生当局より提供された患者の移動情報を、3.1 章前処理で用いた空間分解能、時間分解能に揃え生成した配列を想定している。この患者側の移動情報のフォーマット、ならびに、前処理は、別文書にて定義するものとする。本計算は、配列の同一位置にあるセル ID 間の単純な比較処理であり、どの時刻においても一致していない場合、接触リスクはないものとし、以後の計算より除外する。

なお、感染症患者は、常に 1 名ずつ発生するとは限らず、時には集団発生する。したがって、上記の比較処理も、複数患者を対象に行うケースがありうる。また、接触リスク評価対象とする端末位置と感染者位置とが隣接するセルに位置する場合、両者の実際の位置がセル境界であったとすると接触可能性がゼロとは言えなくなる。その際には、比較処理における一致条件を緩和する必要があるが、実際の処理においては一患者の位置情報配列を複数化(患者位置の上下左右のセル情報を保持した配列を作成)したうえで比較する方が低コストかも知れない。同様に、検索対象端末と患者とが入れ違いで同一エリアに進入したケースを除外してしまわないよう、隣接した時間帯に同一セルに存在しているケースは、接触の可能性ありと判断すべきかも知れない。これは、とりわけ、空気感染する感染症患者が換気の悪い空間にいた際に考慮すべきシナリオと考えられる。

3.3 位置情報を利用した接触リスクの高精度推計

接触リスクの高精度推計に際しては、患者側の位置情報と携帯端末側の位置情報双方の空間的分解能、時間的分解能を上げた処理を行う必要がある。この患者側の位置情報については、「積極的疫学調査」によって、詳細な立ち寄り先や時間、交通手段を収集しうる。それらの情報から正確な座標情報を抽出するには、いくつかの技術を組み合わせる必要がある。たとえば、訪問先として与えられる地名や店舗名から座標へと変換するには、各種 **geocoding** 技術が求められる。また、利用した交通手段情報から座標情報へと変換するためには、**GTFS(General Transit Feed Specification)**[13][14]等を組み合わせる必要がある。これらの詳細については別文書に委ねる。

接触リスクの高精度推計においては、まず、感染者側が滞留していた時間帯と移動していた時間帯とを分離する。患者が移動している際、滞留している端末との接触リスクは低いものとみなし得る。同様に、患者が滞留している際、移動している端末との接触リスクは低いものとみなし得る。各携帯端末が滞留しているか移動しているかは、とある単位時間内に携帯端末の在圏情報が遷移しているか否かにより判断しうる。以下では、「感染者がとある位置に滞留し、端末側も同一地域に滞留中と判断される際の接触リスク」と、「感染者が移動中であり、端末側も移動中である際に生じうる接触リスク」に分け、それぞれの携帯端末側情報の前処理と高精度計算について概説する。

3.3.1 感染者が滞留していた時間帯における接触リスク評価

感染者がいずれかの地域に滞留し、携帯端末側も同一の地域に滞留していた時間帯においては、両者の接触リスクが高まることになる。しかし、携帯基地局のセル面積は数平方キロメートルに及びうるため、同一地域に滞留していることは必ずしも両者の接触を意味しない。提案方式では、端末の所在フロアの区別も行えない。NTT ドコモの分布統計は、端末位置情報を **250m** メッシュへと変換しているが、人口密度が高い首都圏等大都市圏においては同一メッシュ内に存在する端末数は数百～数千へと上りうる。したがって、接触リスク計算にはさらなる高精度が望まれる。

位置情報の高精度化に際しては、携帯の基地局情報には限界がある。その際、**Wi-Fi** 基地局のアクセス履歴を利用することができれば、より高精度な位置情報を取得しうる可能性がある。携帯端末 **GPS** は、一般的に、携帯アプリケーション側が利用するデータであり、携帯キャリア側が利用することは困難である。しかしながら、携帯端末の保有者は、この位置情報を自らの健康管理に役立てる動機を有していることに加えて、アプリケーション事業者側は、収集位置情報の応用が増えることにより収集データの価値を高めることが出来る。したがって、将来的に、携帯キャリアがアプリケーション事業者側へと位置情報を照会する手段が提供されるような展開に至れば、携帯キャリア側における接触リスク計算に **GPS** 状況を活用しうる可能性は存在する。携帯電話端末によっては、**iPhone** のように端末位置情報を一週間程度の記録しているケースもある。これら外部データの活用は、位置情報の保存に求められる携帯キャリア側のコスト軽減にも資するものと考えられる。

3.3.2 感染者が移動していた時間帯における接触リスク評価

次に、感染者が移動していた時間帯における接触リスク評価に際して、感染リスクとしてまず考慮すべきなのは、携帯端末の保有者が、感染者と同じ電車やバス、飛行機等の公共交通機関に同乗していたケースである。この計算において、感染者の移動に関しては、徒歩や自家用車、公共交通機関等、移動手段についての情報が与えられる。一方、携帯端末側については、携帯の在圏情報という粗い粒度の位置情報のみが得られる。以下では、この感染者が移動している時間帯における接触リスク評価について記す。

まず、それぞれの端末における移動速度の概算から、移動手段の種別を推計する。移動速度は、移動元と移動先の基地局とその座標が得られるとの仮定に基づき、基地局間の直線距離を端末のセル移動に要した時間で除すことで概算する。そのうえで、移動手段別の速度の目安(表 3-1)を用いて、可能性のある移動手段を得る。感染者が、徒歩ないし自転車移動中と想定される場合は、公共交通機関のように密閉空間での移動でないことから、感染リスクは低いものとみなす。ただし、渋滞情報を加味した計算が可能な場合、渋滞が発生している地域の低速移動端末は、接触リスクを1段階上げることが考慮される。

表 3-1. 移動手段別の速度の目安

種別	速度	備考
徒歩	4.8km/時(80m/分)	不動産広告で用いる基準値
自転車	15km/時	徒歩のおおよそ3倍
バス	10~20km/時	路線バスの表定速度
	50~70km/時	高速バスの表定速度
自動車	15~60km/時	高速道路以外の法定速度
	80~100km/時	高速道路の法定速度
鉄道	30km/時	都市部の地下鉄の表定速度
	30~90km/時	JR 在来普通列車の表定速度
	50~110km/時	JR 在来特急の表定速度
	120~230km/時	新幹線の表定速度
船舶	52km/時	28ノット(1ノット=1852m/時)

感染者が自動車・バス・鉄道・航空機・船舶等での公共交通機関により移動していることが想定される場合は、同じ交通機関に乗り合わせているかどうかの判定が重要となる。同一交通機関に乗り合わせている端末同士は、一定期間に渡って端末の在圏情報に高い類似性が生じるものと考えられる。そこで、患者の移動時間において、位置情報の交差に関する Jaccard 係数を計算し、一定値以上の類似性を示した端末を、接触リスクがあるものと判定する。感染者とは異なる乗り物に乗っていることが明らかな場合は、接触リスクなしと推定する。

なお、患者側の移動情報は、患者への聞き取りに基づく移動情報となり、具体的な乗車情報が得られているケースが少なからず存在する(バス停〇〇から△△までバス移動、等)。この詳細情報を生かすためには、携帯端末の位置情報を公共交通機関の移動情報にマップする必要がある。一つの手としては、この時点で候補として残っている携帯端末の在圏情報のトレースを対象としてクラスタリングを行ったうえで、それぞれのクラスタを同一公共機関の乗客とみなし、GTFS ([13][14])等の公共交通機関情報[15]、GIS[16]のデータ(バス停、駅、空港、港など)より得られる車両情報と照合する手がある。しかしながら、こうした計算は技術として未確立であり、将来的な検討課題となる。また、患者の移動情報には、詳細な位置情報(たとえば、新幹線〇〇号△△号車××座席等)が含まれるケースがある。この車両内の位置まで加味した接触リスクの計算には、車両内 Wi-Fi 情報等を含むさらに別の技術が必要となる。将来的には、指定席の予約情報など、異業種間のデータの組み合わせによる可能性も考慮する必要がある。

3.4 接触リスク計算に関する検討事項

3.4.1 接触リスクの評価について

接触リスクの概算、接触リスクの高精度推計による評価の結果、本手法による接触リスクはだまかに以下の5段階が得られることになる(表3-2)。このうち、リスク段階のBまでは、在圏情報により認定しうるが、精度的には限界があることに留意する必要がある。リスク段階Aを認定するためには、GPSやWi-Fi基地局情報等を用いた接触リスク評価を行う必要がある。その際は、使用する位置情報によって接触リスクの意味が異なりうる。また、患者側の移動情報も、必ずしも正確ではない場合がありえる。その際には、表3-3に示すとおり、接触リスクの判定を緩やかにする必要がある。

表 3-2. 接触リスクの5段階

リスク	説明	根拠
A	接触リスクが特に高いと判断される状態 (感染者と同一場所、乗り物等に存在した)	空間分解能・時間分解能が高い情報を用いて、接触の可能性が確認された群
B	接触リスクが高いと判断される状態 (感染者と同一場所、乗り物等に存在した可能性がある)	基地局情報を用いた高精度計算にて、接触の可能性がある群
C	基地局情報を用いた計算により、接触リスクがないとは言えない状態 (感染者と同一場所、乗り物等に存在した可能性を否定できない)	基地局情報を用いた高精度計算にて、接触を否定できない群
D	感染者と同一地域にいた可能性がある状態	基地局情報を用いた高精度計算での除外群
E	感染者と交差した可能性が極めて低い状態	基地局情報を用いた概算での除外群

また、接触リスクに加えて、接触のスタイルや病原体の種類に応じて変化しうる「感染リスク」についても考慮する必要がある。たとえば、空気感染する麻疹や結核のように、患者が移動した後にも感染する可能性がある場合も、接触リスクの判定を緩和する必要があることに留意する。狭い区域に滞留している場合でも、屋外の球場や競技場のように換気が良いスペースと屋内の閉鎖空間とでは、感染リスクにも違いが生じる。後者を適切に処理するためには、施設の種類等に関する新たなデータセットを利用する必要がある。一方、前者の処理に際しては、疾患の種類別に応じた接触判定の緩和で対応しうる。

表 3-3. 位置情報の精度の違いによる評価

		評価対象者の位置精度	
		精度低	精度高
感染者の位置精度	精度低	基地局エリアが同一時間帯に重なる場合に接触リスクあり	感染者の基地局エリアに対し、対象者の滞留場所・経路に一定のバッファを設けたものが同一時間帯に重なる場合に接触リスクあり。
	精度高	感染者の滞留場所・経路に対し一定のバッファを設け、それと対象者の基地局エリアが同一時間帯に重なる場合に接触リスクあり。	両者の滞留場所・経路に対し一定のバッファを設け、それらが同一時間帯に重なる場合に接触リスクあり。

3.4.2 位置情報の記憶容量と接触リスクの計算量

携帯端末の位置情報履歴を一定期間残すためには、相応の記憶容量が要求される。接触リスク計算は常時行われるわけではなく、患者が発生する際のみ(当局より患者情報の提供を受けた際)行えば良く、さらに、接触リスク計算には前処理を行う前提となっている。そのため、プライバシー保護と記憶容量の節約を目的として、位置情報履歴はリングバッファ様に管理したうえで、常時圧縮しつつ保存するのが効率的かも知れない。携帯の位置情報は、連長圧縮や差分圧縮に適した特性を持つものと考えられるため、効率的に可逆圧縮しうる可能性がある。

位置情報に基づく接触リスクの計算は、膨大な2次元平面上のユークリッド距離計算となりうるが、利用できる患者側データと携帯端末側データの特性が異なるため、本来的に単純比較ができない。提案手法は、そうした制約を逆手に取り、計算対象の端末数を効率的に削減しうるものと考えられる。次章に示すように、接触リスク計算は顧客からの依頼に基づいて行う整理となっているが、パンデミック時など、社会の関心が高まっている際には多量の依頼が集中する可能性が高い。その際には、リアルタイムな演算ではなく、記録した位置情報全体を対象としたバッチ処理を行うと共に、感染リスクの計算結果を別途キャッシュし、依頼に対して効率的に返答する体制を整える必要があるかも知れない。また、高精度計算を行うための計算資源に比して計算依頼数が多すぎる場合、計算精度を落とすことにより応答性を確保する策も用意しておく必要があるだろう。これは、患者数が増大したタイミングにおいては、現実的に生じるシナリオと考えられる。

3.4.3 プライバシーへの配慮

接触リスクの評価結果のうち、ハイリスク群のリストは、公衆衛生上極めて有意義な情報となる。しかしながら、当データはプライバシーに関わる情報であり、計算基盤より外部に持ち出されるべきでないデータとなる。とりわけ、公衆衛生当局への提供は想定していない。一方で、計算対象となった端末総数が何件で、そのうち、各リスク区分に属する端末の人数等の統計情報は、個々人のプライバシーを毀損する可能性が少なく、今後の感染拡大予測への入力として利用しうる等、公衆衛生上有意義なデータとなりうる。本統計情報については、患者情報を提供する公衆衛生当局に対して、携帯キャリア側より提供する形が望ましい。

なお、携帯位置情報の感染症対策応用に関しては、携帯端末契約者の個人情報保護、通信の秘密の保護に加えて、患者側のプライバシーへの配慮が求められる。感染症患者の情報公開に関しては、感染症法(感染症の予防及び感染症の患者に対する医療に関する法律)上、第十六条に関連規定が存在する(表 3-4)。提案手法は患者情報そのものを公開しないため、同法の立法主旨にも合致したものであるが、その妥当性について別途有識者検討会を設置し検討を進めている。

なお、技術的には、有リスクとされた携帯端末契約者が多数集まり計算結果をシステムティックに統合することで、患者の分布位置を推定できる可能性が存在する。患者プライバシーに対する本リスクに対しても、上記検討会にて検討を進める。

表 3-4. 感染症法

感染症の予防及び感染症の患者に対する医療に関する法律

第十六条 厚生労働大臣及び都道府県知事は、第十二条から前条までの規定により収集した感染症に関する情報について分析を行い、感染症の発生の状況、動向及び原因に関する情報並びに当該感染症の予防及び治療に必要な情報を新聞、放送、インターネットその他適切な方法により積極的に公表しなければならない。

2 前項の情報を公表するに当たっては、個人情報の保護に留意しなければならない。

4. 接触リスクの通知サービス

携帯端末の位置情報を感染症対策に利用することは、携帯位置情報データの目的外利用に当たる。したがって、携帯位置情報を用いて接触リスク計算をする際にも、携帯契約者に対して、プライバシーに関わる携帯位置情報を感染症患者への接触リスク計算に利用する旨を十分に理解して頂いたうえで、明示的な同意を取る必要がある。

この同意取得に際しては、大きく 2 通りの方法が考えうる。まず、携帯キャリアよりユーザーにサービス説明メールを送付し、そのメールを基点として顧客より接触リスク計算の同意と依頼を受ける方法がある。以後、この方式をプル型サービスと称する。プル型サービスでは、感染症患者が発生するたびに、携帯キャリアが同意取得メールを送付することになる。携帯利用者からすると、この動作は煩雑であり、緊急地震速報のように、リスクがある際のみ携帯キャリアより連絡がある形式がより利便性が高い。そこで、患者が発生した際に自動的に接触リスク計算を行って良いか否かの事前同意を取得し、感染リスクがある際のみ通知する方法が考えられる。以後、この方式をプッシュ型サービスと称する。

全ての携帯端末契約者は、初期状態においては感染症患者との接触リスク計算の同意をしていない。したがって、感染者が発生したタイミングで、プル型サービスの案内送付を行うか、国外にて新興感染症等が発生し社会的な関心が高まっているタイミングでプッシュ型サービスの案内送付を行うか、両者のメリットを生かす形で、プル型サービスの案内にプッシュ型サービスの申し込み案内を同報するか、同意取得の経路を多様化しておくことが合理的であろう。

表 4-1. 接触リスクの通知サービスの形態

プル型サービス	携帯電話ユーザーに、サービス説明メールを送付し、その説明に従い、感染者との接触リスク計算に際した位置情報利用への同意とサービス申し込みを受け付ける方式
プッシュ型サービス	感染者との接触リスク計算に際した位置情報利用に対して、携帯電話ユーザーより事前同意を取得し、接触リスクがある場合に通知する方式

4.1 接触リスク通知のプル型サービス

プル型サービス(図 4-1)では、まず、(1) 携帯キャリアがユーザーにサービスの説明メールを送付する。本メールは、感染症患者との接触リスク通知サービスを概説し、サービス利用に際して携帯位置情報を使用することの説明と利用許諾の取得を目的とする。実際には、メールにはサービス概説のみを記し、文中に含まれる URL 等を踏むことにより、同意取得とサービス申し込みページへと遷移する形式も想定される。次に、(2)ユーザーは、自らの携帯位置情報の利用を承諾し、接触リスクの計算を申し込む。これを受けて、携帯キャリアは、第 3 章に示した接触リスク計算を行う。

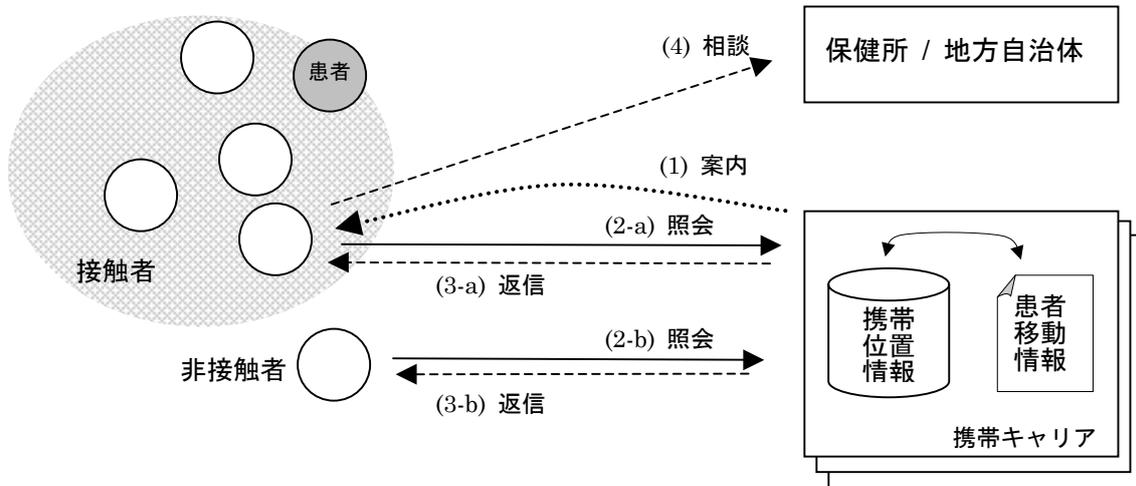


図 4-1. プル型のサービス

そのうえで、携帯キャリアは、(3) 接触リスクの程度をユーザーに通知する。接触リスクが存在することを通知された接触者は、そのリスクの程度により、(4) 最寄の保健所に、感染の可能性について相談することになる。説明メールと結果通知メールの例を表 4-2、表 4-3 に示す。

本モデルでは、計算を要するたびに位置情報の利用許諾を行い、接触リスク評価の申し込みを行う必要がある。この方式は、接触リスクがある場合でもキャリア側からはユーザーにリスクを通知できないデメリットがある反面、ユーザーにとっては、自身の位置情報の利用を限定しうる安心感が得られやすいというメリットがある。

なお、本モデルの発展形として、ユーザーが計算を申し込む際に、スマートフォンアプリ等で収集している詳細な位置情報履歴を携帯キャリア側に提出頂いたうえで、詳細な接触リスク計算に役立てるというアプローチも考えうる。

表 4-2. 接触リスクの通知サービスの案内メール例

利用者各位

感染症等接触リスク通知サービスのご案内

平素は、〇〇〇〇をご愛顧いただき誠に有難うございます。
このたび、感染症等の発生に対し、患者との接触可能性について評価し、結果をお知らせするサービスを行うこととなりましたので、ご案内いたします。

1. 感染症等接触リスク通知サービスの内容

「感染症等接触リスク通知サービス」(以下、「本サービス」といいます)は、〇〇(以下、「当社」といいます)が、厚生労働省、A 社、B 社とともに提供する個人向けのサービス

です。
新型インフルエンザ、結核、麻疹などの感染力の高い感染症が発生した場合に、患者の居住地域や移動経路などの情報をもとに、当該患者との接触の可能性を評価し、お知らせするサービスです。接触の可能性が高い場合には、保健所等の相談窓口についてもあわせてお知らせいたします。

2. お申込みの注意事項

- ・ 本サービスでは、ご利用の携帯電話端末の位置情報(基地局情報、GPS 機能、Wi-Fi 機能の利用によって計測される位置情報)を利用し、接触リスクを計算します。
- ・ リスクの計算に際しては、できる限り精度の高い手法を用いますが、低リスクの判定となった場合でも感染の可能性はゼロではありません。
- ・ 同様に、高リスクの判定となった場合でも、必ず感染が生じるわけではありません。
- ・ お客様の位置情報は接触リスクの判定のみに用い、位置情報、計算結果共に、弊社外の第3者に通知することはありません。

3. お申し込み方法

感染症等接触リスク通知サービス利用規約を御確認頂いた上で、下記リンクをクリックしてください。リスク計算結果は、別途、メールにてお伝え致します。

サービス利用規約: https://*****

サービス申し込み: https://*****

4. 本サービス、ならびに利用者情報の取り扱いに関する問い合わせ窓口

本サービス、ならびに本サービスにおける利用者情報の取り扱いに関して、ご意見、ご要望がございましたら、下記窓口までご連絡くださいますようお願いいたします。

窓口名称: *****

お問い合わせ方法: *****

以上

表 4-3. 接触リスクの通知サービスの結果通知メール例

感染症等接触リスク通知サービス

〇〇様

厚生労働省の令和〇年〇月〇日付け「〇〇の患者の発生について」に関し、〇月〇日時点の当該患者との接触リスク計算結果をお知らせいたします。

(接触リスクがない場合)

・現在感染が確認されている患者の立ち寄り先や移動経路からは、〇〇様の感染患者との接触の可能性は極めて低いと考えられます。
・ただし、今後の感染の拡大によっては、新たに接触リスクが生じる場合があります。下記を参考に、感染防止に努めてください。

(接触リスクが高い場合)

・現在感染が確認されている患者の立ち寄り先や移動経路から、〇〇様の感染患者との接触の可能性があります。

下記を参考に、何らかの症状がある場合には、速やかに下記、〇〇にご連絡ください。

特段の症状がみられない場合には、*****。

(参考)〇〇感染症について

◆どのような病気か

◆どうやって感染するか

4.2 接触リスク通知のプッシュ型サービス

次に、プッシュ型サービスについて、概要を整理する(図 4-2)。プッシュ型サービスでは、まず、(1) 携帯キャリアがユーザーにサービス案内を送付する。ユーザーは、その案内を受けて、(2) サービスに申し込み、携帯位置情報を感染症患者との接触リスク計算に用いることの利用許諾を行う。こうしたサービス案内は、プル型サービスにてリスクの計算結果を通知した際に、プロモーションを兼ねて案内するのも合理的と考えられる。次に、サービス依頼をした携帯端末を対象に、(3) 新たな患者発生の情報が届くたびに、接触リスク計算を行う。その結果は、(4) 一定以上の接触リスクがあると判断された端末を対象として、メール通知する。ユーザーは、その結果を見て、(5) 保健所に相談するか否かを判断する。

プッシュ型サービスでは、事前に携帯位置情報の利用許諾が得られているうえ、接触リスク計算への依頼登録が行われているため、感染症患者の情報提供があればすぐに必要な計算を開始することが出来るメリットがあります。また、感染症が終息するまで登録者の接触リスク計算を継続的に実施でき、感染リスクがある場合には即座に通知できるというメリットがある。事前許諾のメリットとして、普段の通勤経路・勤務先等から感染リスクを推測・予測し、経路の変更や外出禁止等を通知するサービスなどへの発展も考えられる。

デメリットとして、感染症対策に限定されているとはいえ、携帯位置情報が何らかの計算に常時利用され続けられることに対して、ユーザーが不安や懸念を持つ可能性がある。この不安や懸念の払拭のためには、まずはサービスの利便や意義を体感して頂くことが有益と考えられるためプル型サービスの利用を優先し、その利用者に対してサービスの情報提供をするのが合理的かも知れない。

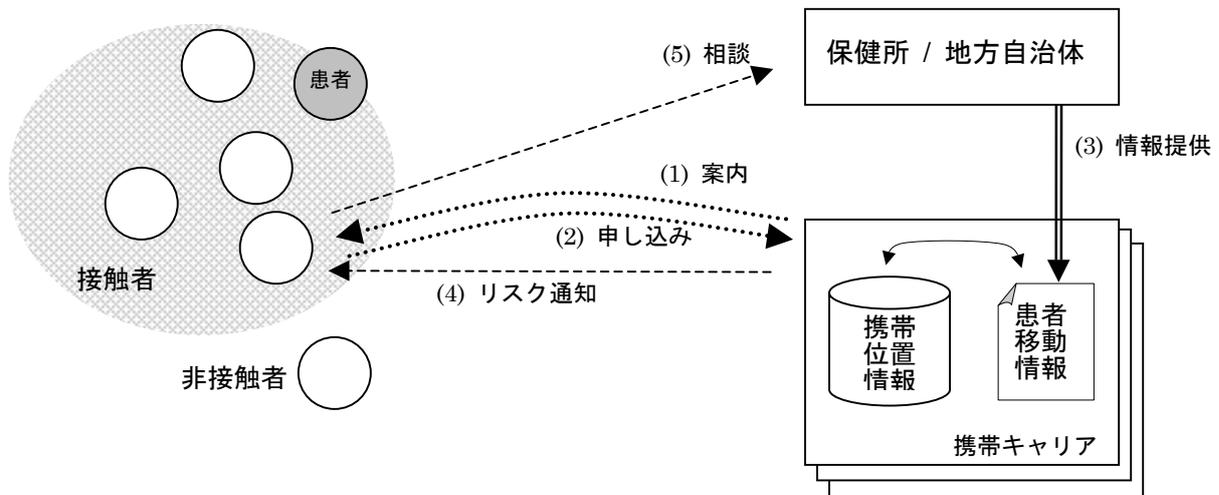


図 4-2. プッシュ型のサービス

4.3 位置情報活用と関係法令

通信事業者が取扱う情報や感染情報は、きわめて機微な情報が含まれ、「通信の秘密」「プライバシー」「個人情報」の観点での整理が必要となる。通信の秘密は、日本国憲法第二十一条に示されており、国民が保有する権利であることがしめされている(表 4-4)。プライバシーについては、日本国憲法第十三条にて、個人として尊重されなければならないことが示されている(表 4-5)。個人情報については、個人情報保護に関する法律において、個人の権利利益の保護が示されている(表 4-6)。以下に、関連箇所を抜粋する。

表 4-4. 日本国憲法 第二十一条

<p>日本国憲法（昭和二十一年憲法）第二十一条を抜粋</p> <p>第二十一条 集会、結社及び言論、出版その他一切の表現の自由は、これを保障する。</p> <p>2 検閲は、これをしてはならない。通信の秘密は、これを侵してはならない。</p>

表 4-5. 日本国憲法 第十三条

<p>日本国憲法（昭和二十一年憲法）第十三条を抜粋</p> <p>第十三条 すべて国民は、個人として尊重される。生命、自由及び幸福追求に対する国民の権利については、公共の福祉に反しない限り、立法その他の国政の上で、最大の尊重を必要とする。</p>

表 4-6 個人情報保護に関する法律

<p>個人情報保護に関する法律（平成十五年法律第五十七号） 第一条及び第二条の一部抜粋</p> <p>第一章 総則 （目的）</p> <p>第一条 この法律は、高度情報通信社会の進展に伴い個人情報の利用が著しく拡大していることに鑑み、個人情報の適正な取扱いに関し、基本理念及び政府による基本方針の作成その他の個人情報の保護に関する施策の基本となる事項を定め、国及び地方公共団体の責務等を明らかにするとともに、個人情報を取り扱う事業者の遵守すべき義務等を定めることにより、個人情報の適正かつ効果的な活用が新たな産業の創出並びに活力ある経済社会及び豊かな国民生活の実現に資するものであることその他の個人情報の有用性に配慮しつつ、個人の権利利益を保護することを目的とする。</p> <p>（定義）</p> <p>第二条 この法律において「個人情報」とは、生存する個人に関する情報であつて、次の各号のいずれかに該当するものをいう。</p> <p>一 当該情報に含まれる氏名、生年月日その他の記述等（文書、図画若しくは電磁的記録（電磁的方式（電子的方式、磁気的方式その他の他人の知覚によっては認識することができない方式をいう。次項第二号において同じ。）で作られる記録をいう。第十八条第二項において同じ。）に記載され、若しくは記録され、又は音声、動作その他の方法を用いて表された一切の事項（個人識別符号を除く。）をいう。以下同じ。）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。）</p>

また、通信事業者は、電気通信事業法において、利用者の権利である「通信の秘密」を保護する責務が示されている(表 4-7)。なお、これらの通信の秘密は、有線・無線にかかわらず適用される。インターネットを通じた有線回線については有線電気通信法(表 4-8)、無線を通じた場合には電波法(表 4-9)が適用され、それぞれの法律にて、電気通信事業法第四条一項の秘密の保護を適用している。

表 4-7. 電気通信事業法

<p>電気通信事業法（昭和五十九年法律第八十六号） 第一条から第四条を抜粋</p> <p>第一章 総則 （目的）</p> <p>第一条 この法律は、電気通信事業の公共性にかんがみ、その運営を適正かつ合理的なものとするとともに、その公正な競争を促進することにより、電気通信役務の円滑な提供を確保するとともにその利用者の利益を保護し、もつて電気通信の健全な発達及び国民の利便の確保を図り、公共の福祉を増進することを目的とする。</p>

<p>(定義)</p> <p>第二条 この法律において、次の各号に掲げる用語の意義は、当該各号に定めるところによる。</p> <p>一 電気通信 有線、無線その他の電磁的方式により、符号、音響又は影像を送り、伝え、又は受けることをいう。</p> <p>二 電気通信設備 電気通信を行うための機械、器具、線路その他の電氣的設備をいう。</p> <p>三 電気通信役務 電気通信設備を用いて他人の通信を媒介し、その他電気通信設備を他人の通信の用に供することをいう。</p> <p>四 電気通信事業 電気通信役務を他人の需要に応ずるために提供する事業（放送法（昭和二十五年法律第百三十二号）第百十八条第一項に規定する放送局設備供給役務に係る事業を除く。）をいう。</p> <p>五 電気通信事業者 電気通信事業を営むことについて、第九条の登録を受けた者及び第十六条第一項の規定による届出をした者をいう。</p> <p>六 電気通信業務 電気通信事業者の行う電気通信役務の提供の業務をいう。</p> <p>(検閲の禁止)</p> <p>第三条 電気通信事業者の取扱中に係る通信は、検閲してはならない。</p> <p>(秘密の保護)</p> <p>第四条 電気通信事業者の取扱中に係る通信の秘密は、侵してはならない。</p> <p>2 電気通信事業に従事する者は、在職中電気通信事業者の取扱中に係る通信に関して知り得た他人の秘密を守らなければならない。その職を退いた後においても、同様とする。</p>

表 4-8. 有線電気通信法

<p>有線電気通信法（昭和二十八年法律第九十六号） 第九条を抜粋</p> <p>(有線電気通信の秘密の保護)</p> <p>第九条 有線電気通信（電気通信事業法第四条第一項又は第百六十四条第三項の通信たるものを除く。）の秘密は、侵してはならない。</p>

表 4-9.電波法

<p>電波法（昭和二十五年法律第百三十一号） 第五十九条を抜粋</p> <p>(秘密の保護)</p> <p>第五十九条 何人も法律に別段の定めがある場合を除くほか、特定の相手方に対して行われる無線通信（電気通信事業法第四条第一項又は第百六十四条第三項の通信であるものを除く。第百九条並びに第百九条の二第二項及び第三項において同じ。）を傍受してその存在若しくは内容を漏らし、又はこれを窃用してはならない。</p>

通信事業者が取扱う情報は、きわめて機微な情報が含まれることから「通信の秘密」「個人情報」「プライバシー」の関係法令を遵守することが求められる。これらの関係法令は、個人に権利が有すること、および、みだりに個人の権利を侵すことを禁止することを基本的な考え方としている。しかし、これらの利用にあたって、利用目的・方法・公開範囲等などの明確な境界はなく、それぞれの立場によって見解が分かれることもあるが、事前承諾を得ている場合はこの限りではない。事前承諾とは、前章までに記載するように、利用者からの申し込み等により、予め「通信の秘密」「個人情報」「プライバシー」の利用の承認を得ることである。

利用者からの承諾の方法、または、承諾または申し込みの時期(タイミング)は、前節までに示した、サービスの提供方法、通知方法や運用方法に大きく関係する。事後承諾とならないように配慮・留意した運用が必要となる。また、利用目的や方法・公開の範囲をきわめて限定させ、目的外の利用や個人の不利益が生じないための仕組みづくりも必要となる。

5. 接触リスク計算の外部委託について

高精度な接触リスクの計算のためには、対象とする位置情報として、基地局情報以外に Wi-Fi や GPS 等の情報を利用する必要が生じる。これらの情報は、携帯キャリアが保有している情報だけでなく、アプリベンダーや携帯製造元等が保有する情報[17]が含まれる。したがって、リスク評価の高精度化に向けては、携帯キャリアが保有する位置情報だけでなく、3.3.1 において言及したように、位置データを保有する組織への効率的な照会や、データ提供の枠組みを整える必要が生じる。

また、接触リスク計算サービスの運用に際して、携帯キャリアは、以下のようなインフラ整備とサービス体制の構築を行う必要がある。構築・運用に求められるコストの多くは、新たな位置情報の保存のためのインフラと考えられるが、それ以外の体制については、各キャリアがサービスのためのインフラを構築し、維持することはコスト面で非効率である可能性がゼロではない。

- ・ 接触リスク計算を行うためのサーバ等の機器の設置とその維持管理
- ・ 感染症患者発生時の接触リスク計算
- ・ 保健所/自治体から提供される患者情報(患者属性、患者移動情報)の保有・管理
- ・ 患者情報取り扱い部門に対する教育訓練、流出防止対策
- ・ 患者情報と契約者情報のマッチング
- ・ 電車、バス等の情報と合わせて計算を行う場合には、それらのデータのセットアップ・更新

そこで、将来的な可能性として、「ユーザーからの依頼に基づいて、ユーザーに関係する位置情報を集約し、リスク計算を行う第3者機関」のような仕組みが必要となるかも知れない(図 5-1)。この場合、接触リスクの計算を行う第3者機関は、患者情報を提供する保健所/地方自治体との間で秘密保持契約を締結し患者情報の提供を受けるとともに、携帯キャリア等との間でも秘密保持契約を締結し、位置情報等の提供を受けて感染リスク計算を行う。この方法の問題点は、各携帯キャリア等の保有する契約者の位置情報を、第三者機関に受け渡す必要があることである。契約者情報の受け渡しリスクを最小限にする方法としては、携帯キャリア等において契約者 ID や携帯電話番号などの個人情報とダイレクトに紐づけられる情報を除外するなど、一定レベルの秘匿化をしたデータに加工した上で、接触リスク計算実施団体にデータを受け渡す方法が考えられる。

ただし、ユーザーによる承諾と依頼によるとは言え、キャリアの保有する位置情報を組織外に提供することには大きなリスクが伴う。そのリスクを避けるためには、各社が保有する位置情報は各社に留め置きながら、各社における計算結果を統合してリスク評価を行うような形式も考えられる。その場合の接触リスク通知サービスの構成イメージは、図 5-2 の通りとなる。

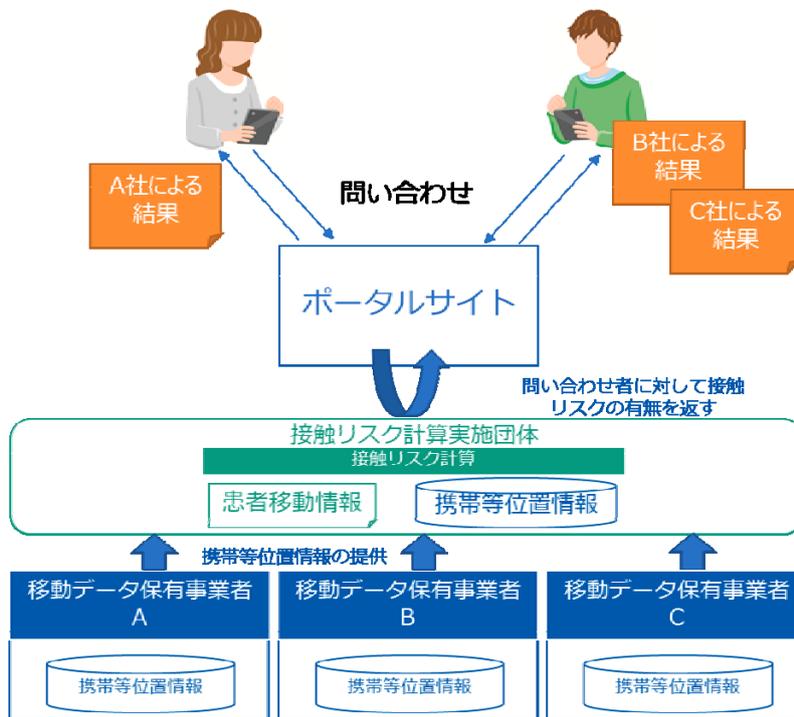


図 5-1. 接触リスク計算を行う第 3 者機関の利用イメージ

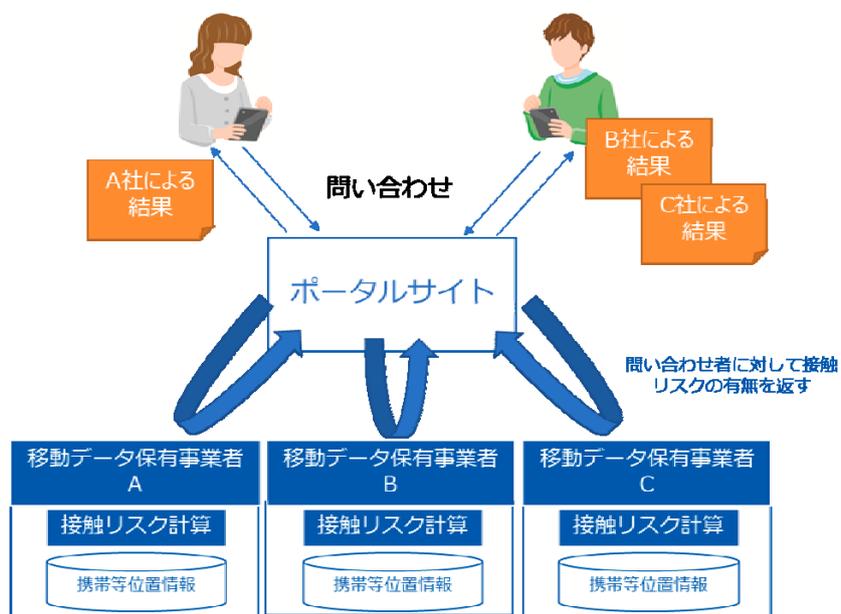


図 5-2. 第三者機関で接触リスク計算を統合する場合の構成イメージ

ここまで示した接触リスク計算通知サービスを、表 5-1 に整理する。このような接触リスク計算の外部委託先としては、実際に商用サービス等でデータ処理を実施している企業(2.4 に挙げた企業)が考えられる。ただし、その構築、運営には多額のコストを要することに加えて、コスト負担のモデルを検討する必要がある。

表 5-1. 接触リスク計算通知サービス

	サービスの概要	接触リスク 計算実施者	携帯キャリアによる 情報の保有		携帯等位置情報 の取り扱い
			患者情報	接触者情報	
方式1	携帯キャリア自身が接触リスク計算を行う	携帯キャリア	保有	保有	携帯キャリアから外部には出ない
方式2	接触リスク計算は第三者機関が実施する	第三者機関	保有しない	保有しない	第三者機関に提供する
方式3	携帯キャリアの接触リスク計算を第三者機関が統合	携帯キャリア	保有	保有	携帯キャリアから外部には出ない

参考文献

- [1] T. Okumura, "Tracing infectious agents with mobile location information: A simple and effective countermeasure against epidemic risks", 2019 IEEE Global Humanitarian Technology Conference (GHTC 2019), Seattle, October 2019.
- [2] 総務省 緊急時等における位置情報の取扱いに関する検討会 報告書「位置情報プライバシーレポート ～位置情報に関するプライバシーの適切な保護と社会的利活用の両立に向けて～」平成 26 年 7 月。
https://www.soumu.go.jp/menu_news/s-news/01kiban08_02000144.html
- [3] 鈴木啓介ほか「LTEを収容するコアネットワーク(EPC)の開発」 NTT DOCOMO テクニカル・ジャーナル Vol.19 No.1.
- [4] 電子情報学会 知識ベース 知識の森 4 群 3 編 2 章 2-3「移動通信のコアネットワーク」 藪崎正実。
http://www.ieice-hbkb.org/files/04/04gun_03hen_02.pdf#page=15
- [5] IJ フォーカス・リサーチ(1)「フル MVNO とは何か、IJ はなぜフル MVNO を目指すのか」Internet Infrastructure Review (IIR) Vol.38, 2018 年 3 月 30 日。
<https://www.ij.ad.jp/dev/report/iir/038/02.html>
- [6] 3GPP TS 29.272 v15.10.0(2019-12) 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Evolved Packet System (EPS); Mobility Management Entity (MME) and Serving GPRS Support Node (SGSN) related interfaces based on Diameter protocol (Release 15).
- [7] 「LTE のさらなる発展-LTE Release9-」 NTT DOCOMO テクニカル・ジャーナル、VOL.18 NO.1, 2010. https://www.nttdocomo.co.jp/corporate/technology/rd/technical_journal/bn/vol18_1/048.html
- [8] Survey of Cellular Mobile Radio Localization Methods: from 1G to 5G
http://spcomnav.uab.es/docs/journals/DelPeral_Survey_preprint_2017.pdf
- [9] 株式会社 NTTドコモ モバイル空間統計 <https://mobaku.jp/>
- [10] 株式会社 NTTドコモ「基地局を利用した位置情報提供インタフェース仕様書第 1.1 版」(2011 年 7 月 20 日)
https://www.nttdocomo.co.jp/binary/pdf/service/developer/smart_phone/base_station/location/location_interface_v1.1.pdf
- [11] 「携帯電話ネットワークからの統計情報を活用した社会・産業の発展支援—モバイル空間統計の概要」 NTT DOCOMO テクニカル・ジャーナル, Vol.20, No.3, pp.6-10,2012.
https://www.nttdocomo.co.jp/corporate/technology/rd/technical_journal/bn/vol20_3/
- [12] 株式会社ワイヤ・アンド・ワイヤレス(Wi2)
<https://wi2.co.jp/jp/privacy-utilization/>
- [13] バス情報フォーマット(GTFS リアルタイム)ガイドライン初版(2019 年 3 月 27 日)
<http://www.mlit.go.jp/common/001283242.pdf>
- [14] 静的バス情報フォーマット(GTFS-JP)仕様書(第2版)(2019 年 3 月 27 日)
<http://www.mlit.go.jp/common/001283244.pdf>
- [15] 公共交通オープンデータセンター, <https://www.odpt.org/>, <https://ckan.odpt.org/dataset>
- [16] 国土交通省「国土数値情報ダウンロードサービス」, <http://nlftp.mlit.go.jp/index.html>
- [17] Apple「位置情報データに関する Apple の Q&A」2011 年 4 月 27 日
<https://www.apple.com/jp/newsroom/2011/04/27Apple-Q-A-on-Location-Data/>

Tracing infectious agents with mobile location information

A simple and effective countermeasure against epidemic risks

Takashi Okumura

Kitami Institute of Technology
Kitami, Hokkaido, Japan
National Institute of Public Health, Japan
Wako City, Saitama, Japan
Email: tokumura@mail.kitami-it.ac.jp

Abstract—The Sustainable Development Goals set by the United Nations request all countries to strengthen the capacity for early warning, risk reduction and management of national and global health risks. This study aims to reduce the risks caused by infectious agents by identifying individuals who accidentally come in close contact with affected patients, utilizing mobile location information. The proposed approach involves asking public health authorities to provide the trip data of patients to mobile carriers under a non-disclosure agreement. The carriers then determine the possible risks of contact of their mobile customers with patients and disclose only the risk information to their customers upon requests. This approach provides novel channels for public health authorities to announce the risk of contagious agents to residents, while maintaining the privacy of both the patients and residents. For a nationwide deployment of the proposed approach, several issues are discussed in this study, along with brief summaries of related studies.

Index Terms—Location information, Mobile devices, Public Health, Infection control

I. INTRODUCTION

Technological developments in the 20th century have greatly reduced the risk of infectious diseases for mankind via vaccinations, antibiotics, and antivirals. Nevertheless, infectious diseases still remain a major public health concern in the 21st century. Infectious diseases spread through droplets, direct contacts, and airborne agents. Among these diseases, tuberculosis, chicken pox, and measles, which are highly contagious, are transmitted by air. Such diseases can exponentially spread to people as one affected patient can easily infect multiple people. Accordingly, public health authorities need to prevent the spread of infection. The Sustainable Development Goals set by the United Nations [1] also request all countries to strengthen the capacity for early warning, risk reduction and management of health risks. To this end, it is necessary to first identify patient's cohabiting family and close contacts sharing room with patients on a daily basis at work or in school and to carefully monitor their signs and symptoms so that prompt responses can be provided when needed.

Meanwhile, the more difficult task is to identify third parties who accidentally come into contact with the patient, e.g., individuals who are in continuous contact with patients in a poorly ventilated space, such as public transportation (e.g., deadly packed commuter trains in Tokyo metropolitan area). In Japan, public health authorities announce the patient's residence area and trip information to inquire about potential

individuals who come in contact with the patient for reporting to public health centers. However, such information does not interest many residents unless they have particular concern regarding the case. The actual address of the patients and their travel details may help in more efficiently alerting the residents. However, such information is highly private, and public health authorities are unable to disclose this in detail. Consequently, authorities can only announce partial information regarding the patient such as data regarding large-scale commercial facilities they visit and public transportation they use, where the patient may come in potential contact with healthy individuals. As stated, public health authorities are unable to effectively provide such information to healthy individuals who may come in potential contacts with patients when identification measures are needed to prevent further infections or to avoid an epidemic.

In Japan, the possession rate of mobile phones exceeds 100% [2]; thus, the location information of these devices may be exploited for public health operations. The straightforward approach is to use device locations to filter people who have no possible contact with patients by determining the distance between the patient in question and other individuals. However, such location information is highly private; thus, government use of such information must be strictly limited for specific cases such as criminal investigations.

In this study, we addressed this problem by devising a way to realize effective infection control while protecting the privacy of patients and residents. This proposal is still under investigation, and there are challenges for its nationwide deployment by the authorities. However, the proposed method can be helpful for other countries and regions experiencing similar problems. The following section illustrates the design of the proposed simple solution and discusses issues involved with comments on related works, followed by concluding remarks.

II. PROPOSED DESIGN

Public health authorities cannot publicly disclose patient's trip data for privacy reasons. In the proposed approach, public health authorities disclose patient's residential and trip data only to mobile carriers who have signed a non-disclosure agreement. Under this agreement, the mobile carrier cannot disclose patient information by itself to any third party.

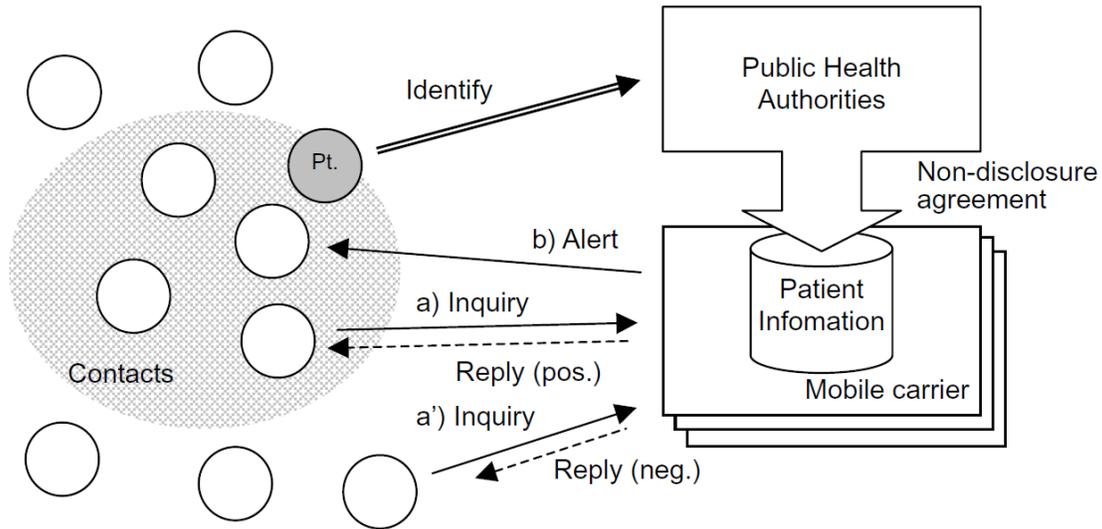


Fig. 1. Overview of the proposed approach

Now, mobile carriers can calculate the distance and contact time between the patients and their customer's devices using the location information they have and the trip data of the patient provided. The carrier may send messages to customers asking whether they require risk information or not. Then, upon request by their customers, they only reply the possibility of contacts (Fig. 1-a). If the risk is low, responses to the requesting customers would indicate that the risk is low (Fig. 1-a'). If a customer registers for the risk notification service in advance, the mobile carrier may send such alerts when the possibility of contact is high (Fig. 1-b).

In the messages that are sent to such customers who may come in possible contact with patients, the alerts may contain guidance to contact public health centers available nearby. The message may also include links to web systems that interactively collect necessary information on the behalf of public health centers and provide necessary instructions, thereby providing effective and efficient responses to public health threat.

In this proposed model, patient information remains private and protected. The information regarding residents is not sent to public health authorities unless they intentionally provide such information. Consequently, this approach effectively and efficiently contributes to public health while protecting the privacy of both patients and residents.

III. DISCUSSION

The utilization of mobile location information in public health has been mostly reported in the context of health promotion, particularly through Pokémon GO. There have been social requirements to motivate people to go outside and improve movement. In such cases, the game was found to be a new model to promote healthy actions [3], [4]. Mobile games have been reported to have both positive (higher activity) and negative (physical hazards) effects on public health [5], [6].

Nevertheless, it is notable that the game demonstrated the value of using location information in public health.

Another application of mobile location information for public health was demonstrated in 2015 during the outbreak of Middle East respiratory syndrome (MERS) in South Korea [7]. Public health authorities tracked the locations of mobile phones and credit card bills to trace patients [8]. As a countermeasure against extreme threat to public health, such as in the cases of MERS and Ebola hemorrhagic fever, patient tracking is justified for public health protection. However, it raises concerns regarding the fact that the government has excessively intervened in public privacy, and such an application must be restricted to the bare minimum necessary requirement.

Our proposed approach helps in providing an effective response for controlling infections while protecting the privacy of both the patients and residents. Currently, Japan is experiencing a steady increase in the number of foreign visitors who travel throughout Japan. Thus, there is a need for countermeasures against imported infectious diseases. This need is imminent, particularly for the upcoming Tokyo Olympic Games in 2020. Japan has the highest population density among the developed countries and thus has the highest risk of epidemics, which necessitates effective countermeasures against infectious agents. The proposed method has desirable properties for handling this type of public health threat.

Notably, several technical options are available to locate mobile devices. First, the mobile carriers possess base station information that each device is connected to. It is easy to convert the base station location into longitude and latitude. However, the granularity of the locational information gained through this approach is coarse; thus, this approach could provide an unnecessarily large number of potential contacts. There are several possible solutions to this problem. First, the wave strength of base stations could be used for positioning

of the device. Second, the information on Global Positioning System (GPS) can be made available on certain occasions. Third, information on Wi-Fi connections can be also used. Further, this information can be combined with geographic and public transportation databases, to correct the positions for obtaining more accurate estimation of individuals who are at a risk of contact with patients.

IV. CONCLUDING REMARKS

This manuscript proposes an application of mobile location information for public health, which provides novel channels for public health authorities to deliver the risk of contagious agents to residents. The proposed approach enables an efficient and effective announcement of such risks while maintaining the privacy of patients and residents.

Several issues remain to be solved before the practical deployment of the proposed method. First, it is necessary to establish societal consensus by clarifying ethical issues involved in the proposal. Second, there must be legal and organizational foundations for actual operation by public health authorities. In Japan, there are laws related to infection control and crisis management; however, there must be a justification from the authorities that give patient information to private companies under non-disclosure agreements. Similarly, it will be necessary to define procedures needed to deliver private information across organizations. Third, there exist technical issues in the implementation of this method. For example, it is necessary to prove that the calculation of contact risk is feasible even with 100 million mobile devices.

We envisage that the proposed approach is deployed as a simple and effective countermeasure against possible epidemic risks, throughout Japan. To this end, we are planning to put the approach to trial, during the Tokyo Olympic Games of 2020, where millions of foreign guests travel throughout Japan. It would be a desirable opportunity to demonstrate the potential of the approach, and the experiment would then contribute for the early warning, risk reduction and management of health risks caused by contagious agents worldwide.

REFERENCES

- [1] United Nations, "Transforming our world: The 2030 agenda for sustainable development," *General Assembly 70 session*, 2015.
- [2] Ministry of Internal Affairs and Communications, "Quarterly on the number and share of telecommunications service contracts," http://www.soumu.go.jp/menu_news/s-news/01kiban04_02000136.html, 2018.
- [3] A. G. LeBlanc and J.-P. Chaput, "Pokémon Go: A game changer for the physical inactivity crisis?" *Preventive medicine*, vol. 101, pp. 235–237, 2017.
- [4] M. Tateno, N. Skokauskas, T. A. Kato, A. R. Teo, and A. P. Guerrero, "New game software (Pokémon Go) may help youth with severe social withdrawal, hikikomori," *Psychiatry research*, vol. 246, p. 848, 2016.
- [5] C.-C. Yang and D. Liu, "Motives matter: motives for playing Pokémon Go and implications for well-being," *Cyberpsychology, Behavior, and Social Networking*, vol. 20, no. 1, pp. 52–57, 2017.
- [6] V. R. Wagner-Greene, A. J. Wotring, T. Castor, J. Kruger, and S. Morte-more, "Pokémon GO: Healthy or harmful?" *American journal of public health*, vol. 107, no. 1, p. 35, 2017.
- [7] K. Kim, T. Tandi, J. Choi, J. Moon, and M. Kim, "Middle East respiratory syndrome coronavirus (MERS-CoV) outbreak in South Korea, 2015: epidemiology, characteristics and public health implications," *Journal of Hospital Infection*, vol. 95, no. 2, pp. 207–213, 2017.

- [8] C. Lee and M. Ki, "Strengthening epidemiologic investigation of infectious diseases in Korea: lessons from the middle east respiratory syndrome outbreak," *Epidemiology and health*, vol. 37, 2015.

TRACING PATIENT PLOD BY MOBILE PHONES

MITIGATION OF EPIDEMIC RISKS BASED ON PATIENT LOCATIONAL OPEN DATA

Ikki Ohmukai

Graduate School of Humanities and Sociology
The University of Tokyo, Tokyo, Japan
i2k@l.u-tokyo.ac.jp

Yasunori Yamamoto

Database Center for Life Science
Research Organization of Information and Systems, Tokyo, Japan
yy@dbcls.rois.ac.jp

Maori Ito

PLOD info, Japan
maorinphone@plod.info

Takashi Okumura

National Institute of Public Health, Japan
Kitami Institute of Technology, Kitami, Hokkaido, Japan
tokumura@mail.kitami-it.ac.jp

March 16, 2020

ABSTRACT

In the cases when public health authorities confirm a patient with highly contagious disease, they release the summaries about patient locations and travel information. However, due to privacy concerns, these releases do not include the detailed data and typically comprise the information only about commercial facilities and public transportation used by the patients. We addressed this problem and proposed to release the patient location data as open data represented in a structured form of the information described in press releases. Therefore, residents would be able to use these data for automated estimation of the potential risks of contacts combined with the location information stored in their mobile phones. This paper proposes the design of the open data based on Resource Description Framework (RDF), and performs a preliminary evaluation of the first draft of the specification followed by a discussion on possible future directions.

1 Introduction

In the cases when public health authorities confirm a patient with highly contagious disease, such as measles, tuberculosis, and the novel coronavirus, they have to identify close contacts of infected patients to monitor their health conditions for a certain period [1].

This *contact tracing* is mostly performed manually by officials. However, concerning the usage of public transportation and visiting public facilities, the officials are unable to trace all possible contacts. To mitigate the potential risks associated with such situations, public health authorities release the reports on the travel information of the infected patients, expecting potential contacts to report themselves. However, due to privacy concerns, authorities cannot provide the detailed information about the patients, and therefore, these releases typically comprise the data related to commercial facilities and public transportation used by the patients.

Most residents rarely pay attention to such releases, except the families who have members in immunocompromised conditions, for example. Newspapers and TV programs might broadcast the releases; however, such news usually have a limited number of readers or audiences, and therefore, reach only a small part of the population.

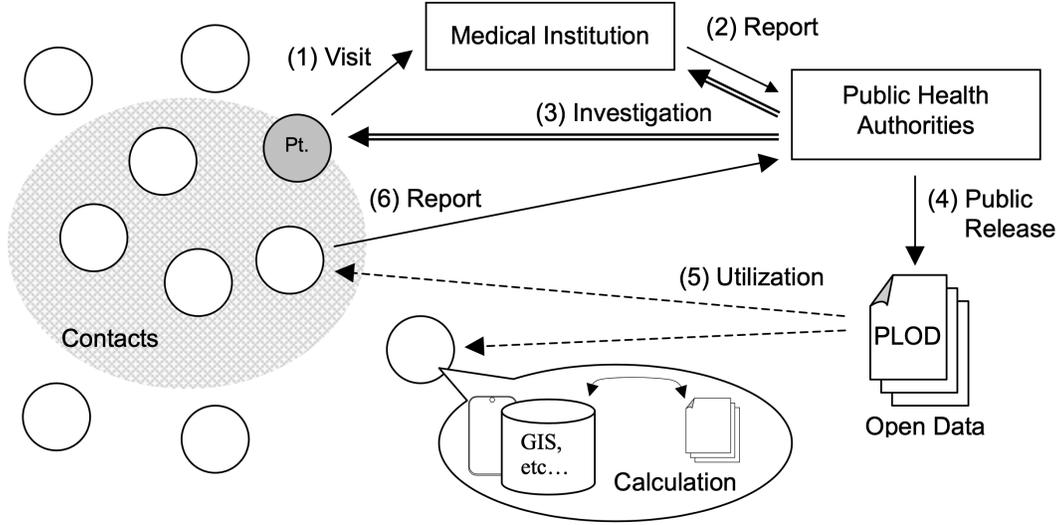


Figure 1: Overview of the proposed approach

Accordingly, public health authorities are not capable of efficiently alerting the potential contacts about the risk associated with the infectious agents, who are mostly unconcerned. We consider that sharing the open data on the patient location information can be used to efficiently bridge this gap by *personalizing the alerts for each resident through the automated estimation of potential intersections between the resident and the infected patients* based on the location history stored in their own mobile phones. Such approach can be helpful, particularly, when patients have visited various places, and when the number of patients has increases to a certain threshold exceeding the capacity of ordinary people to process the trace information corresponding to all patients.

The proposed open data framework for tracing the patient information is illustrated in Figure 1. First, a patient visits a medical institution and a physician confirms the infection case (1). Secondly, the physician reports to a regional public health center (2). Third, the officials perform the necessary investigation (3) and publicly alert the information by an announcement (4). In the proposed scheme, this announcement is released as the open data established in a form of the Patient Locational Open Data (PLOD). These data can be publicly used to perform mapping of patient traces and estimate the risks of potential contacts for individuals (5). To perform this estimation, residents can employ the location history information on their mobile phones. A suspicious contact is supposed to be reported to the public health authority with the purpose of obtaining further instructions (6).

The rest of the paper is organized as follows. First, Section 2 provides an overview on the related research focused on the use of the mobile location information in the public health field. Section 3 outlines the design of the proposed open data scheme. Section 4 demonstrates the preliminary evaluation of the proposed approach, followed by Section 5 to discuss the advantages and limitations. Section 6 concludes the paper.

2 Related works

Utilization of the mobile location information for the public purposes was initiated in the late 2000s, when mobile phones became commonly available across the world [2]. It was deemed as a straightforward approach to trace patient by analyzing the location information provided by mobile devices. Such application was first realized as a countermeasure for the epidemic of Middle East Respiratory Syndrome (MERS) in 2015, reportedly [3]. However, the trace information is considered as the privacy data, which is highly sensitive concerning individuals, and therefore, there are ethical concerns even considering public health applications [4]. Due to the privacy issue, it has been hardly possible to utilize the personal information to facilitate public health preparedness, and only a limited number of related research works have been published.

At the same time, the research works on the statistical utilization of the mobile location information have gained wide popularity, as mobile phones produced a vast amount of the location information relevant for the improvement of population health [5, 6, 7, 8]. In this regard, the statistical utilization of the mobile location information to control infectious diseases has been considered as a promising approach [9, 10, 11, 12]. This field of research has been investigated when the location data related to phone calls were used for efficient surveillance of post-disaster situations

and for managing the quarantine of cholera [13, 14, 15], which was then applied to modeling the spread of infectious diseases [16].

It should be noted that these studies utilized the caller-receiver information of mobile devices referred to as Call Detail Records (CDR). In the related studies, researchers basically used the anonymized or statistically processed data. Nevertheless, there were concerns about privacy issues even in these cases, and the related studies were mostly conducted in developing countries where the associated regulations were less strict. In such countries, mobile phones became common owing to their efficiency as social infrastructure means, and healthcare applications based on such devices also became popular, for example, such as m-health [17]. However, in developing countries, the majority of population use prepaid devices, and therefore, mobile carriers do not possess subscriber contracts. Accordingly, the trace data are not accompanied with the demographic information, which is a severe limitation associated with this approach [12].

In 2020, the situation has drastically changed due to the pandemic of the new coronavirus infectious disease (COVID-19). In China, the government authority have deployed the facial recognition technology, security cameras, and social media analysis combined with the manual surveillance to automate the contact tracing with regard to COVID-19 patients [18]. For this purpose, the government have collected necessary records of calls and the location information provided by transportation and communication companies, which are owned by the government. In South Korea, public health authorities have monitored the credit card history, security camera footage, location information obtained from mobile devices, public transportation cards, immigration records, etc., to track individuals exposed to the risks of possible infection [19]. In Taiwan, the subjects of the home quarantine have been monitored by the authority, reportedly, through analyzing mobile phone signals [19]. Such applications of processing the location information with the purpose of crisis management of infectious diseases have been rarely published as research articles before the novel coronavirus pandemic in 2020, except for the cases of patient tracing in South Korea [3] and a proposal on a risk assessment service of potential infections [20].

The challenge associated with the knowledge representation of the information about infected patients and their location data lies in the fact that such approach should represent both the qualitative information corresponding to attributes and conditions of patients and quantitative information on the location, such as coordinates, distance, and range. Conventionally, these data need to be handled by authorized independent systems. In this regard, analyzing the geospatial information has gained less attention in the information retrieval field. Even concerning the systems with an extension to perform geospatial estimations, the interoperability between systems is not guaranteed [21]. To address the problem, the World Wide Web Consortium (W3C) proposed a framework that can consistently handle qualitative and quantitative information in terms of geospatial knowledge called GeoSPARQL [22]. GeoSPARQL has been implemented as an extension to the resource description framework (RDF) [23] and SPARQL [24], a query language available in RDF, which was standardized by the W3C. GeoSPARQL can be used to realize the representation and retrieval of information in terms of both qualitative and quantitative data and to enable inference operations on the knowledge, providing the foundation for facilitating the representation of patient tracing.

3 Design

In this section, we propose a structured representation of the press releases appropriate to publish the information as PLOD. The press releases representing the current situation about infected patients usually comprise the following information formulated in natural language, which is not suitable for automated processing.

- Issuer and contact information of the document;
- Patient's attributes (age and gender);
- Name of the disease (infected agents) and infectivity;
- Other medical background;
- Clinical course of the patient;
- Travel information.

The metadata include the name of the government agency or local government, as a publisher of the releases, date and time of the publication. The patient information includes demographic attributes such as age, gender, occupation, and place of residence, as well as the information on the disease, such as the infectious agents, and the degree of infectivity. Other medical background and clinical course include the date of onset, confirmed date, and past medical history. Outline of the travel history might be included here as well. The releases contain the information to provide an overview on the patient situation; however, the amount of the information included in a particular release may vary. It may describe a single case or represent only the number of patients infected by a certain disease.

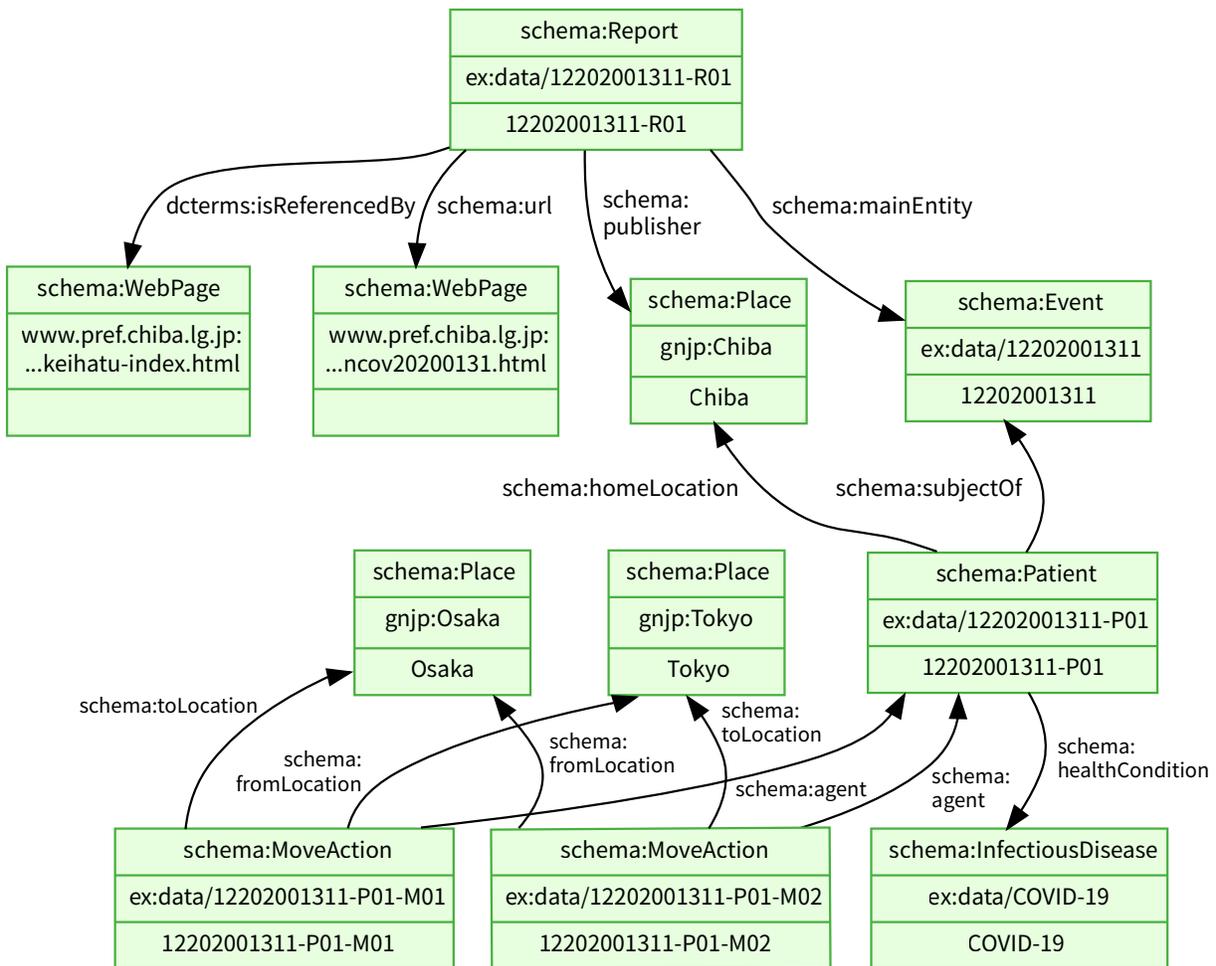


Figure 2: RDF model of PLOD

The travel information of patients is mostly composed of a combination of the following items: i) Travel time or range; ii-a) Name of the facility, or address visited by the patient; or ii-b) The list of public transportation used. In some cases, it may come in a negation form due to denying a visit of a specific place. Figure 2 represents an example of a press release that outlines the two travels of a patient: from Tokyo to Osaka, and from Osaka to Tokyo, as described in the PLOD model.

To model the information, unique and persistent identifiers (PIDs) are required. In the proposed model, each entity has a persistent ID. An event entity is automatically generated based on the press release ID so that all elements are connected through the event entity, even when it is not explicitly mentioned in the press release. Patient IDs are constructed based on the event ID adding a unique number of the patient. To identify the travel information, their IDs are generated by adding an incremental number in the travel history associated with a patient ID. The PLOD URIs are determined uniformly from these IDs by using the HTTP scheme (<https://>), domain name (plod.info), and path information ([/data/](https://plod.info/data/)) as a prefix.

Most of the properties corresponding to each entity employ Schema.org. Missing attributes in Schema.org are originally defined with the prefix "ex:". We select the following properties to represent the patient travel information that is the most important part of the PLOD: "schema:startTime", "schema:endTime", "schema:fromLocation", and "schema:toLocation" defined in the class "schema:MoveAction". The authenticity is ensured by specifying a URI of the original file issued by the local public health authority. Figure 3 represents the raw data of the PLOD model using the turtle format.

```

1 @prefix rdf: <\protect\vrule width0pt\protect\href{http://www.w3.org/1999/02/22-rdf-syntax-ns
#}{http://www.w3.org/1999/02/22-rdf-syntax-ns#}> .
2 @prefix rdfs: <\protect\vrule width0pt\protect\href{http://www.w3.org/2000/01/rdf-schema#}{
http://www.w3.org/2000/01/rdf-schema#}> .
3 @prefix schema: <https://schema.org/> .
4 @prefix dcterms: <\protect\vrule width0pt\protect\href{http://purl.org/dc/terms/}{http://purl.
org/dc/terms/}> .
5 @prefix foaf: <\protect\vrule width0pt\protect\href{http://xmlns.com/foaf/0.1/}{http://xmlns.
com/foaf/0.1/}> .
6 @prefix gnjp: <\protect\vrule width0pt\protect\href{http://geonames.jp/resource/}{http://
geonames.jp/resource/}> .
7 @prefix plod: <https://plod.info/property/> .
8
9 <https://plod.info/data/12202001311> a schema:Event ;
10   rdfs:label "12202001311" .
11
12 <https://plod.info/data/12202001311-R01> a schema:Report ;
13   rdfs:label "12202001311-R01" ;
14   schema:mainEntity <https://plod.info/data/12202001311> ;
15   plod:numberOfPatients "1"^^schema:Integer ;
16   schema:datePublished "2020-01-31"^^schema:DateTime ;
17   schema:publisher gnjp:Chiba ;
18   schema:url <https://www.pref.chiba.lg.jp/shippei/press/2019/ncov20200131.html>;
19   dcterms:isReferencedBy <https://www.pref.chiba.lg.jp/shippei/kansenshou/keihatu-index.html
>.
20
21 <https://plod.info/data/12202001311-P01> a schema:Patient ;
22   rdfs:label "12202001311-P01" ;
23   schema:subjectOf <https://plod.info/data/12202001311> ;
24   schema:healthCondition <https://plod.info/entity/COVID-19> ;
25   plod:dateConfirmed "2020-01-31"^^schema:DateTime ;
26   foaf:age "20s" ;
27   schema:gender "Female" ;
28   schema:homeLocation gnjp:Chiba .
29
30 <https://plod.info/data/12202001311-P01-M01> a schema:MoveAction ;
31   rdfs:label "12202001311-P01-M01" ;
32   schema:agent <https://plod.info/data/12202001311-P01> ;
33   schema:startTime "2020-01-16"^^schema:DateTime ;
34   schema:endTime "2020-01-16"^^schema:DateTime ;
35   schema:fromLocation gnjp:Tokyo ;
36   schema:toLocation gnjp:Osaka ;
37   schema:instrument "Airplane"@ja .
38
39 <https://plod.info/data/12202001311-P01-M02> a schema:MoveAction ;
40   rdfs:label "12202001311-P01-M02" ;
41   schema:agent <https://plod.info/data/12202001311-P01> ;
42   schema:startTime "2020-01-22"^^schema:DateTime ;
43   schema:endTime "2020-01-22"^^schema:DateTime ;
44   schema:fromLocation gnjp:Osaka ;
45   schema:toLocation gnjp:Tokyo ;
46   schema:instrument "Bus"@ja .
47
48 <\protect\vrule width0pt\protect\href{http://geonames.jp/resource/Tokyo}{http://geonames.jp/
resource/Tokyo}> a schema:Place ;
49   rdfs:label "Tokyo" .
50
51 <\protect\vrule width0pt\protect\href{http://geonames.jp/resource/Osaka}{http://geonames.jp/
resource/Osaka}> a schema:Place ;
52   rdfs:label "Osaka" .
53
54 <\protect\vrule width0pt\protect\href{http://geonames.jp/resource/Chiba}{http://geonames.jp/
resource/Chiba}> a schema:Place ;
55   rdfs:label "Chiba" .
56
57 <https://plod.info/entity/COVID-19> a schema:InfectiousDisease ;
58   rdfs:label "COVID-19" ;
59   schema:name "2019-nCoV acute respiratory disease"@en ;
60   schema:infectiousAgent "2019-nCoV" ;
61   schema:code <\protect\vrule width0pt\protect\href{http://purl.bioontology.org/ontology/
ICD10/U07.1}{http://purl.bioontology.org/ontology/ICD10/U07.1}> .
62
63 <\protect\vrule width0pt\protect\href{http://purl.bioontology.org/ontology/ICD10/U07.1}{http
://purl.bioontology.org/ontology/ICD10/U07.1}> a schema:MedicalCode ;
64   schema:codeValue "U07.1" ;
65   schema:codingSystem "ICD-10" .

```

Figure 3: Example of PI OD (excerpt)

Table 1: Press-release of the infected cases published by local governments in Japan

	2018	2019	2020	Total
Detailed description of trace	3	67	1	71
Mentions of trace	1	3	5	9
Non-specific description of trace	30	20	30	80
No public exposure	1	15	0	16
Not available	16	26	3	45
Others	0	3	0	3
No case description	3	1	0	4
Yearly total	54	135	39	228

4 Analysis

4.1 Analysis of press releases for infectious diseases

To verify the appropriateness of the patient information described in the PLOD model, we collected the press releases corresponding to infectious disease patients published by the local governments in Japan. The survey results included the press releases of 189 measles cases and 39 COVID-19 ones starting from 2018 through 2020, in total, 228 cases published by the Ministry of Health, Labor, and Welfare, as well as prefectures and cities. We then categorized the documents according to the way used to describe the case trace information. The obtained results are summarized in Table 1.

Among the 228 releases, 71 described the trace in the independent sections, while 9 releases mentioned the trace in the bodies of the announcements. Another 80 releases included the trace information using generic terms and avoided providing the specific location. Then, 16 releases notified that there were no risks of public exposure by the cases, and 45 releases did not provide any detailed information of the case. In sum, among of 228 releases, 224 contained the case descriptions and 80 included the trace information. The format of the press releases and the contents were not uniform: some of them included the names of specific places, and others did not present the information sufficiently detailed to trace the patients.

Names of the hospitals visited by the patients were in some cases omitted in the announcements. Medical institutions kept the patient records as per each day, and public health authorities and administrators of the institutions could directly contact the possible contacts. Accordingly, in Japan, such information could have been excluded from the releases intentionally, for privacy reasons.

The survey is incomplete in the sense that the analysis has been conducted only based on the releases found by using search engines, and there are releases that can be found only in archives of the local governments. Nevertheless, it is highly probable that the press releases corresponding to the patients infected by contagious agents do not have a single standardized format. In this regard, announcements by public health authorities in English-speaking countries are deemed to have similar characteristics, concluding from the twelve announcements we found by using search engines. The tentative conclusion can be summarized as follows: the information provided by public health authorities is not standardized, and it is considered meaningful to propose a framework for the unified representation of the actual information about infected patients.

4.2 Evaluation of the proposed model

Next, we evaluate the expressive power of the proposed PLOD model based on the press releases that contain the detailed trace information about patients. For the evaluation, we analyzed 39 press releases issued in the period from Jan 16, 2020 to Feb 11, 2020 and arranged the available information according to the PLOD model by using RDF (in the turtle format). The metadata of the press release itself, the patient information, and the patient trace information were successfully expressed for all cases by using PIDs issued for each entity considered as a Subject. However, we confirmed that the information corresponding to Objects became empty for the cases in which the detailed data were not provided in the corresponding release. Among the considered 39 press releases, 23 included the patient information, and 8 out of those 23 represented the trace information of patients, including date and location.

```

1 <https://plod.info/data/29202001281-P01-M01> a schema:MoveAction .
2   schema:agent <https://plod.info/data/29202001281-P01> ;
3   schema:startTime "2020-01-08"^^schema:DateTime ;
4   schema:endTime "2020-01-11"^^schema:DateTime ;
5   schema:fromLocation "" ;
6   schema:toLocation "" ;
7   schema:instrument "Bus" .

```

Figure 4: Example of a failure case

To estimate the contact risk using the PLOD model, it is essential to include the origin of a trace or a destination, preferably, together with the detailed information about the visited places. To obtain the detailed description of the trace information, we extracted the steps from each case report, and 21 steps were identified on the basis of the 8 releases. Among these steps, 10 included the origin or destination data with varying granularity: 8 were described at the prefecture level, 1 as “oversea”, and only 1 contained the facility-level specific information that could be converted into the coordinates (latitude and longitude).

We identified the four patterns in the trace information that could not be expressed appropriately in PLOD. First, the releases included a negation of a specific area, such as “overseas”. Second, several included cases of the profession information, such as “a bus guide” and “a driver”, indicated the possibility of trips but did not specify it in a manner sufficient to trace the case. Figure 4 illustrates a case corresponding to this category. Third, there were the releases that indicated traces in generic terms, for example, “returning home with a child”, which could not be converted into the coordinates. Lastly, several releases referred to the other press releases that could contain the trace information, such as “same as past case A”. It is hardly possible to convert the first and the second cases into the coordinates, whereas the third and the fourth cases might be processed to obtain the required data. For example, we may infer that the “home” is supposed to be in the city where the press release has been issued, and, similarly, the reference in the fourth case might be appropriately processed. Both cases can be organized as the best fit for the SPARQL-based queries that can infer the information even if it is not explicitly stated.

Limitations of this evaluation are associated with the sampling bias related to the fact that the analysis has been conducted only on the limited number of samples without the possibility to apply random sampling. Additionally, the coding has been implemented by a single person, and the statistical analysis is not possible. Nevertheless, the collected samples can be used to illustrate that the proposed framework is capable of expressing the patient location data described in the current press releases, although there are possibilities for the further improvement.

5 Discussion

Concerning the actual deployment of the proposed public health service in Japan, we need to execute several steps as follows. In the Japanese health care system, each local government is responsible for control of infectious diseases, and therefore, public releases are issued by them either in HTML, or PDF format at their websites. In this setting, there are mainly two ways to provide mobile applications with access to the open data. The one approach is to establish a centralized model, in which a national public health authority, such as Ministry of Health, would perform the entire processing, such as collection of releases, conversion of formats, and distribution. The other approach is to implement a decentralized approach in which each local government would issue its announcement in the RDF format using the standardized vocabularies. In the latter case, there would be a need for a public RDF data repository providing mobile applications with the data they need. As both approaches necessitate the substantial coordination between organizations, a reasonable option would be to collect, convert, and distribute the open data via a third party organization, considering that the releases are publicly usable.

An indispensable part of the proposed approach is to develop the mobile applications that would use the open data to raise customized alerts for users. Mobile phones typically have a built-in application programming interfaces (API) for location services, which responds the position data of the mobile phone. Accordingly, mobile applications may periodically call the API to record their positions, which can be later used to estimate the intersections with the patients mentioned in the open data. Once the application finds that a patient enters in a predefined range corresponding to the phone, it may alert the owner of the phone.

At the moment, the proposed framework is at a draft stage, and there could be steps required to implement before the public use becomes possible. As mentioned above, public releases have a variety of formats, and they often contain the

irregular data, which are hardly possible to express in the current version of PLOD. To address the problem, continuous improvement for the specification is necessary, however, the formats of press releases need to be also standardized and updated. To facilitate the process of creating a unified data representation, it would be preferable to launch the service, even with the limited quantity in terms of releases and the limited quality in terms of expressive power. Then, during the actual use, the specification and application would evolve to meet the real needs.

The proposed application can be implemented in several ways. It could be developed as an independent application or as an embedded module in an application that uses a location service, such as applications for crisis management provided by local governments. As the proposed framework is based on the RDF, users or third parties can easily add the information relevant according to their needs. Organizing a hackathon is an approach to develop a variety of applications suitable for various usage options. As the public health information concerning infectious diseases may attract attention of a considerable number of people, the proposed application can be widely installed, once the application becomes available for the public use.

Under the pandemic conditions, there could be a surge of data requests, once the applications are widely installed in the population. Accordingly, the service needs to be scalable aiming to assure the stability of the service as a crisis management system. In this regard, the open data might be distributed via Content Delivery Networks (CDNs). Even a state-of-art RDF store may not be suitable to handle the surge of SPARQL queries from a large number of mobile phones. To mitigate this problem, we may also define a simplified query, equivalent to a full-fledged SPARQL query. SPARQL provides a flexible framework to retrieve various data at the cost of the overhead to process the complex query on the server side. Accordingly, instead of issuing a SPARQL query, each application may issue a simplified queries consisting of a couple of arguments so as to accommodate a higher number of requests, while providing the flexibility of SPARQL.

6 Conclusion

In this study, we have proposed a framework to represent information press releases focused on the current situation on infected patients in a form of the open data using RDF. To the best of our knowledge, this is the first attempt of applying the RDF and open data technology to mitigation of epidemic risks. Due to the ongoing pandemic caused by COVID-19, we consider that the proposed framework has a potential to be used worldwide.

The future research work will include the following three steps. First, we plan to collect the public releases, convert them into the open RDF data format, and distribute them aiming to facilitate the development of applications. This could serve as a demonstration for stakeholders to evaluate the utility of the open public health data concept. Second, the proposed framework needs to be supported by the authorities, such as the ministry of health or a national institution, in a centralized manner. Third, local governments are expected to initiate providing the data in the linked open data format directly in a decentralized manner.

In the field of epidemiology, public health fields have been using information technologies mostly for the purposes of statistics and prediction; however, they have rarely considered using such methods for prevention of diseases. Moreover, the concept of open data in the medical field has not been considered widely, and there exist only a limited number of publication dedicated to this subject. We envisage that the proposed approach may serve as a foundation for development of knowledge processing applications in public health, such as alerting about hospital overload under pandemic conditions.

Acknowledgment

The authors thank Ms. Rie Nariko and Ms. Nobuko Nakagomi for their contribution in the collection of press releases.

References

- [1] S. A. Rasmussen and R. A. Goodman, *The CDC Field Epidemiology Manual*. Oxford University Press, 2018.
- [2] C. Yang, J. Yang, X. Luo, and P. Gong, "Use of mobile phones in an emergency reporting system for infectious disease surveillance after the sichuan earthquake in china," *Bulletin of the World Health Organization*, vol. 87, pp. 619–623, 2009.
- [3] K. Kim, T. Tandi, J. W. Choi, J. Moon, and M. Kim, "Middle east respiratory syndrome coronavirus (mers-cov) outbreak in south korea, 2015: epidemiology, characteristics and public health implications," *Journal of Hospital Infection*, vol. 95, no. 2, pp. 207–213, 2017.

- [4] K. H. Jones, H. Daniels, S. Heys, and D. V. Ford, "Toward an ethically founded framework for the use of mobile phone call detail records in health research," *JMIR mHealth and uHealth*, vol. 7, no. 3, p. e11969, 2019.
- [5] B. Ouedraogo, J. Gaudart, and J.-C. Dufour, "How does the cellular phone help in epidemiological surveillance? a review of the scientific literature," *Informatics for Health and Social Care*, vol. 44, no. 1, pp. 12–30, 2019.
- [6] S. Lai, A. Farnham, N. W. Ruktanonchai, and A. J. Tatem, "Measuring mobility, disease connectivity and individual risk: a review of using mobile phone data and mhealth for travel medicine," *Journal of travel medicine*, vol. 26, no. 3, p. taz019, 2019.
- [7] K. H. Jones, H. Daniels, S. Heys, and D. V. Ford, "Challenges and potential opportunities of mobile phone call detail records in health research," *JMIR mHealth and uHealth*, vol. 6, no. 7, p. e161, 2018.
- [8] E. Jahani, P. Sundsøy, J. Bjelland, L. Bengtsson, Y.-A. de Montjoye *et al.*, "Improving official statistics in emerging markets using machine learning and mobile phone data," *EPJ Data Science*, vol. 6, no. 1, p. 3, 2017.
- [9] J. Chirombo, P. Diggle, D. Terlouw, and J. Read, "A review of models of human mobility for predicting infectious disease spread," *Modelling spatial processes of infectious diseases*, p. 60, 2018.
- [10] C. Panigutti, M. Tizzoni, P. Bajardi, Z. Smoreda, and V. Colizza, "Assessing the use of mobile phone data to describe recurrent mobility patterns in spatial epidemic models," *Royal Society open science*, vol. 4, no. 5, p. 160950, 2017.
- [11] K. Sallah, R. Giorgi, L. Bengtsson, X. Lu, E. Wetter, P. Adrien, S. Rebaudet, R. Piarroux, and J. Gaudart, "Mathematical models for predicting human mobility in the context of infectious disease spread: introducing the impedance model," *International journal of health geographics*, vol. 16, no. 1, p. 42, 2017.
- [12] A. Wesolowski, C. O. Buckee, K. Engø-Monsen, and C. J. E. Metcalf, "Connecting mobility to infectious diseases: the promise and limits of mobile phone data," *The Journal of infectious diseases*, vol. 214, no. suppl_4, pp. S414–S420, 2016.
- [13] L. Bengtsson, X. Lu, A. Thorson, R. Garfield, and J. Von Schreeb, "Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in haiti," *PLoS medicine*, vol. 8, no. 8, 2011.
- [14] F. Finger, T. Genolet, L. Mari, G. C. de Magny, N. M. Manga, A. Rinaldo, and E. Bertuzzo, "Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks," *Proceedings of the National Academy of Sciences*, vol. 113, no. 23, pp. 6421–6426, 2016.
- [15] J. Cinnamon, S. K. Jones, and W. N. Adger, "Evidence and future potential of mobile phone data for disease disaster management," *Geoforum*, vol. 75, pp. 253–264, 2016.
- [16] L. Bengtsson, J. Gaudart, X. Lu, S. Moore, E. Wetter, K. Sallah, S. Rebaudet, and R. Piarroux, "Using mobile phone data to predict the spatial spread of cholera," *Scientific reports*, vol. 5, p. 8923, 2015.
- [17] M.-P. Gagnon, P. Ngangue, J. Payne-Gagnon, and M. Desmartis, "m-health adoption by healthcare professionals: a systematic review," *Journal of the American Medical Informatics Association*, vol. 23, no. 1, pp. 212–220, 2016.
- [18] L. Lin, "China Marshals Its Surveillance Powers Against Coronavirus," <https://www.wsj.com/articles/china-marshals-the-power-of-its-surveillance-state-in-fight-against-coronavirus-11580831633>.
- [19] E.-Y. Jeong, "South Korea Tracks Virus Patients' Travels—and Publishes Them Online," <https://jp.wsj.com/articles/SB12291155354026644516304586207690851702666>.
- [20] T. Okumura, "Tracing infectious agents with mobile location information: A simple and effective countermeasure against epidemic risks," in *2019 IEEE Global Humanitarian Technology Conference (GHTC 2019)*. IEEE, 2019.
- [21] C. Torniai, S. Battle, and S. Cayzer, "Sharing, discovering and browsing photo collections through rdf geo-metadata," in *SWAP 2006 - Semantic Web Applications and Perspectives, Proceedings of the 3rd Italian Semantic Web Workshop*, ser. CEUR Workshop Proceedings, vol. 201. CEUR-WS.org, 2006. [Online]. Available: <http://ceur-ws.org/Vol-201/07.pdf>
- [22] R. Battle and D. Kolas, "Geosparql: enabling a geospatial semantic web," *Semantic Web Journal*, vol. 3, no. 4, pp. 355–370, 2011.
- [23] D. Beckett and B. McBride, "Rdf/xml syntax specification (revised)," *W3C recommendation*, vol. 10, no. 2.3, 2004.
- [24] E. Prud'hommeaux and A. Seaborne, "SPARQL Query Language for RDF," W3C Recommendation, 2008. [Online]. Available: <http://www.w3.org/TR/rdf-sparql-query/>

TRACING PATIENT PLOD BY MOBILE PHONES

MITIGATION OF EPIDEMIC RISKS BASED ON PATIENT LOCATIONAL OPEN DATA

Ikki Ohmukai

Graduate School of Humanities and Sociology
The University of Tokyo, Tokyo, Japan
i2k@l.u-tokyo.ac.jp

Yasunori Yamamoto

Database Center for Life Science
Research Organization of Information and Systems, Tokyo, Japan
yy@dbcls.rois.ac.jp

Maori Ito

PLOD info, Japan
maorinphone@plod.info

Takashi Okumura

National Institute of Public Health, Japan
Kitami Institute of Technology, Kitami, Hokkaido, Japan
tokumura@mail.kitami-it.ac.jp

March 16, 2020

ABSTRACT

感染力の強い感染症に罹患した患者が生じた際、公衆衛生当局は、患者の健康状態や移動先等の情報を開示し、社会に注意喚起することがある。しかし、患者プライバシーの観点から、こうした情報開示においては詳細情報は省かれており、大規模商業施設や公共交通機関等、不特定多数の第三者との接触が生じた可能性に関する必要最小限の情報が留められている。多くの住民はこうした情報に注意を払わないことから、公衆衛生当局は、感染の危険について効果的な情報提供を行うことに困難があった。本研究では、この患者の移動情報を構造化し、オープンデータとして公開するための枠組みを提案する。これにより、住民は、患者の移動情報と携帯電話端末に蓄積される自らの移動情報とを用いて、患者との接触リスクを自動的に計算し、認知することが可能となる。本稿では、この患者情報に関するオープンデータをRDF (Resource Description Framework) にて実現する手法を提案すると共に、日本において公開されている患者情報のプレスリリースを用いて評価し、今後の発展について考察する。

1 緒言

結核や麻疹、新型インフルエンザに代表される感染力の強い感染症が発生した場合、公衆衛生当局は、濃厚接触者を特定して健康監視する必要がある。この接触者調査は、公衆衛生当局による聞き取りにより行われる。しかし、公共交通機関や何らかの大規模施設を訪問した際には、全ての接触者を追跡することが困難となる。こうした不特定多数との接触が生じた際、当局は、患者の移動情報を公開し、接触可能性のある住民に保健所への連絡を呼びかける。しかし、感染症患者の住所や移動経路の詳細情報はプライバシーに深く関わることから、行政機関として詳細に公表することが出来ない。結果として、行政による情報公開は、接触者数が多い大規模商業施設や公共交通機関等の利用情報のみに留まってきた。

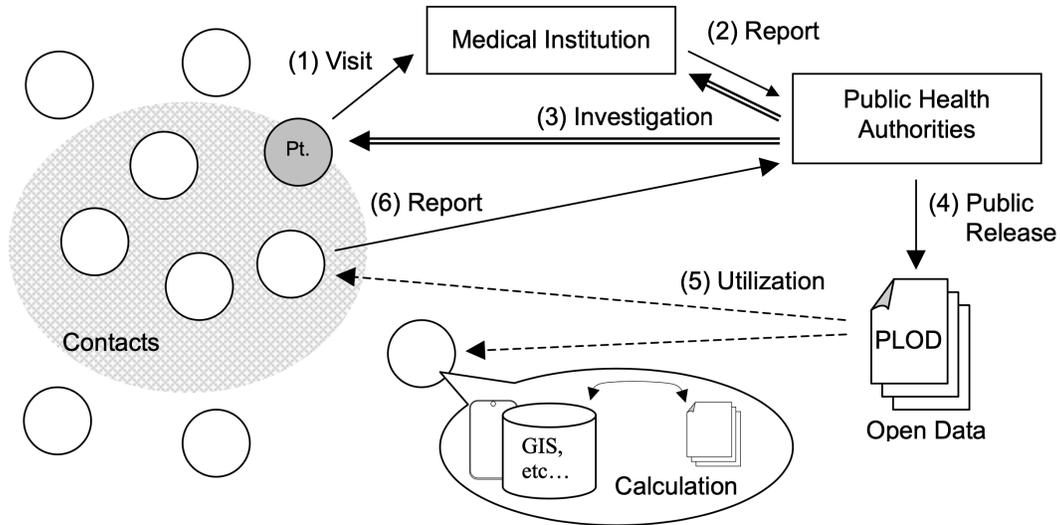


Figure 1: Overview of the proposed approach

住民の多くは、家族に免疫不全患者がいる等のケースを除いて、こうした当局からの情報提供に興味を払わない。プレスに投げ込んだ情報は報道記事にはなるが、そうした記事も拡散力が乏しく、有効な対策に繋がらない懸念があった。このように、感染症の罹患リスクについて、公衆衛生当局は潜在的な接触者に効果的に情報提供することが出来ない。もし、このプレスリリース自体をオープンデータとして公開することができれば、汚染地域の住民は、携帯電話に蓄積される自らの位置情報履歴との交差を計算し、患者との感染リスクを簡単に把握することが可能となる。こうしたアプリケーションは、患者の訪問先が多いケースや、同時発生する患者数が多いケース等、すべての患者情報を検証していくことが困難となっていく場合において、とりわけ有効に機能することが期待される。

そこで本稿では、公開可能な情報として公開された患者の移動情報を対象として、患者の移動情報を「オープンデータ化」するための手法について提案する(図1)。まず、患者が医療機関を受診し、医師が患者を診断する(1)。次に、医師は、患者の発生を保健所に届け出る(2)。さらに、保健所は、患者への聞き取り調査を実施し(3)、患者発生のプレスリリースを公開する(4)。我々は、このプレスリリースを、患者位置情報に関するオープンデータ (Patient Locational Open Data, PLOD) として、規定のフォーマットにおいて公開することを提案する。こうしたデータは、自由に利用することができ、患者移動のマッピングや個人にとっての患者との接触リスク推定に利用することができる(5)。この推定においては、住民側は、自らの携帯電話を用いて記録されていく位置情報を利用する。そのうえで、接触が疑われる個人は、保健所に届け出て、その後の必要な指示を受けることになる(6)。

本稿では、まず、2章において公衆衛生における携帯位置情報活用に関する関連研究を整理する。3章において、提案デザインを概説し、4章において評価を試みる。5章では、実用化に向けた展望と限界について記し、6章に結語を示す。

2 感染症対策と携帯位置情報

携帯電話の位置情報を公益活用するという試みは、携帯電話が普及した2000年代の後半頃に起源がある[1]。もっともシンプルな形として、患者の追跡のために患者携帯の位置情報を利用するというものがあり、2015年に韓国で生じた中東呼吸器症候群 (Middle East Respiratory Syndrome, MERS) 騒動において利用された報告がある[2]。しかしながら、携帯位置情報はプライバシー上極めてセンシティブであり、公衆衛生への活用においても倫理的な議論があった[3]。とりわけ、患者個々の位置情報を扱うことはハードルが高く、研究論文として報告される例も限定されてきた。

一方で、携帯位置情報の統計的な活用に向けては、今まで公衆衛生が得られなかった情報が大量に活用できることから、研究が活発化している[4, 5, 6, 7]。そのなかでも、携帯位置情報の感染症対策への活用が期待されている[8, 9, 10, 11]。この分野は、災害後の効率的な現況調査をコレラ防疫に活用した事例[12, 13, 14]から始まり、携帯位置情報を用いた感染モデルの構築に活用[15]されるようになった。

これらの研究の多くは、Call Detail Records (CDR) と称される通話データによる位置情報を利用している。また、データは基本的に匿名化されているか、あるいは、統計処理したデータを用いる形となっている。それでも、プライバシー問題への懸念があるようで、関連論文はそうした制約が緩和される発展途上国における研究が多い。インフラ整備の効率性から携帯電話が普及している発展途上国においては、「m-health」として知られる、携帯電話の医療・公衆衛生応用が盛んとなっていたことも背景にあるものと考えられる[16]。ただし、発展途上国では、プリペイド式の携帯電話の利用が多く、携帯キャリアも端末の保有者情報を保有していない。そのために、各位置情報も、保有者の属性情報と紐付けられておらず、研究上の制約となることが指摘されてきた[11]。

さて、こうした状況は、2019年より生じた新型コロナウイルス感染症 (COVID-19) のパンデミックによって、大きく変わりつつある。中国では、顔認証やセキュリティーカメラ、ソーシャルメディア管理を通常の人間による監視と組み合わせ、COVID-19患者の接触者追跡を自動化したと言われている[17]。そのために、携帯電話の通話情報、位置情報など各種サービスの利用記録を企業から収集したとされる。中国においては、国内の交通・通信網を管理する企業の多くが国営であることから、迅速にデータを活用することが可能であった。韓国では、保健当局が、クレジットカードの使用記録、防犯カメラ映像、携帯電話の位置情報、公共交通用カード、入国管理記録などを精査することが可能であり、感染者や感染リスクのある人の移動履歴を突き止めている[18]。台湾でも、自宅隔離の対象者が携帯電話の位置情報を利用して追跡されているとされる[18]。こうしたサービスは、韓国における患者追跡の事例[2]と筆者らによるわが国における患者との接触リスク通知サービスの提案[19]を除いて、新型コロナウイルス感染症の大流行前には研究論文としての公開はほとんど知られていない。

患者とその移動に関する知識表現の難しさは、個人の属性などの定性的情報と、距離や範囲などの数値的な位置情報を混在させる必要が生じる点にある。従来は異なる検索手法に対してはそれぞれに適したシステムを構築して結果を統合する必要があった。また、知識検索の研究分野において、そもそも位置に関する情報検索があまり注目されてこなかった事情もある。検索エンジンが独自の拡張によって位置情報の検索に対応している事例もあるが、相互運用性が担保されないという問題もある[20]。そこで、定性的な情報と定量的な情報の双方を1つのデータベースに格納し、同一クエリ中で記号と位置情報の指定や検索を行うことができるGeoSPARQLが、World Wide Web Consortium (W3C) より提案された[21]。GeoSPARQLは、W3Cにより標準化されたResource Description Framework (RDF) [22]ならびにクエリ言語SPARQL[23]の拡張として定義されており、データの表現方法、取得方法のいずれについても相互運用性が確保されている。そのため、GeoSPARQLは、患者に関する情報の知識表現に用いることができることに加えて、定性的、定量的なデータを対象とした柔軟な情報検索ニーズに応えることができる。こうした長を備えたGeoSPARQLを活用することで、PLODは、患者の追跡情報を表現し、また、推論処理を行っていくうえで、基盤となる特徴を兼ね備えることになる。

3 PLODモデルの設計

この章では、患者情報をPLODとして公開するうえで、プレスリリースに含まれる情報の構造化について提案する。現在、日本においてプレスリリースを通じて公開される感染者に関する情報には、以下のような情報が自由記述されており、自動処理を行うことが難しい。

- 文書の発行者・連絡先情報等
- 患者自身の属性情報 (年齢帯や性別)
- 患者の罹患している感染症情報 (名称や感染力の程度)
- その他医学的背景情報
- 患者自体の大まかな臨床的経過
- 患者の移動情報

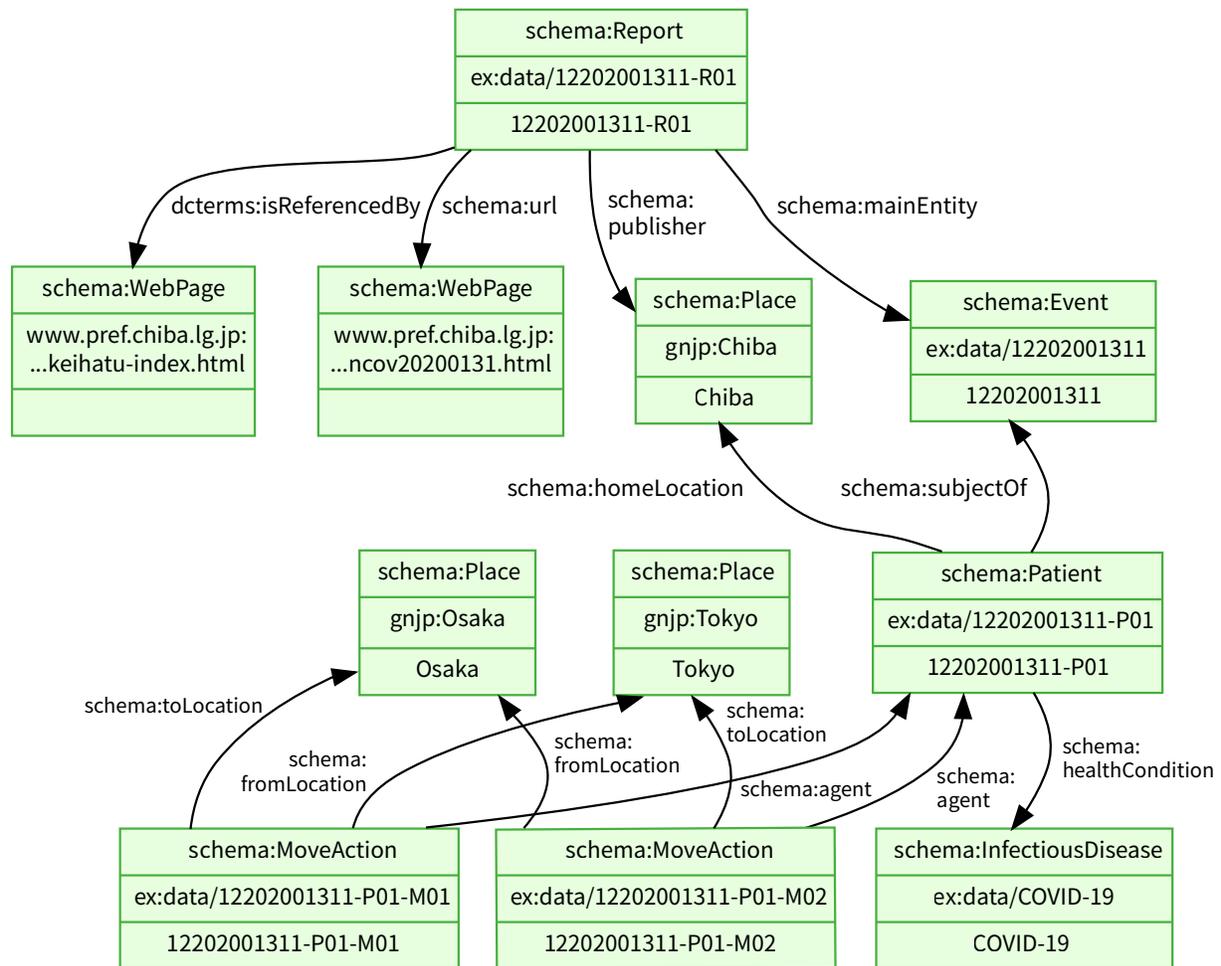


Figure 2: RDF model of PLOD

メタデータにはプレスリリースの発信者である政府機関、あるいは自治体の名称、公表日時が含まれる。患者情報には、患者情報にはデモグラフィック属性として年齢、性別、職種、居住が含まれ、疾患情報として感染症名や感染力の程度等が含まれる。その他の医学的背景情報として、発熱日、確定日、臨床定期経過、既往歴など、今回の感染症に関わる患者の医学的情報が公開される。そのうえで、移動情報の概要や渡航歴が記載される。移動情報には日時あるいは日時の範囲、場所が含まれる。特定の場所に立ち寄っていないなどの否定的情報についてもここに記述される。このように、プレスリリースには、患者の状況が概観できる状況が含まれる。ただし、実際に提供される情報の質や量は、プレスリリースごとに大きく変わる。プレスリリースによっては、とある患者について詳述することもあるが、とある疾患に何名が罹患したかという患者数のみの公開に留まることもある。

患者の移動情報の多くは、下記の要素の組み合わせで表現される。i) 移動した時間、ないし、時間帯; ii-a) 患者が訪問した施設や住所; ないし ii-b) 利用した公共交通機関のリスト、である。ただし、時には、とある場所への訪問を否定する形の情報となっている場合もある。プレスリリースに含まれる患者の移動情報をPLODモデルにて表現したものを図2に示す。この例では、東京から大阪に移動した後、大阪から東京に戻ったケースを示している。

情報をモデルで表現していくためには、永続的なID (PIDs) が求められる。現在のモデルでは、各エンティティにこの永続的なIDが付与されている。モデル化するにあたっては、明示的には記載されていな

Table 1: Press-release of the infected cases published by local governments in Japan

	2018	2019	2020	Total
Detailed description of trace	3	67	1	71
Mentions of trace	1	3	5	9
Non-specific description of trace	30	20	30	80
No public exposure	1	15	0	16
Not available	16	26	3	45
Others	0	3	0	3
No case description	3	1	0	4
Yearly total	54	135	39	228

イベントエンティティを仮定し、このイベントエンティティを中心にそれぞれの要素を接続する。イベントエンティティのIDは、プレスリリースのIDから機械的に生成される。患者のIDは、イベントエンティティのIDにプレスリリース中の患者の通し番号を加えたもので構成する。移動情報については、同一人物の移動履歴にインクリメンタルな番号を与えた上で患者のIDと連結させたものをIDとする。PLODのURIは、これらのIDに接頭辞として一律にHTTPスキーム (<https://>) とドメイン名 (plod.info)、パス情報 ([/data/](#)) を付加したものとす。

各エンティティのプロパティは、Schema.orgで提供されるプロパティを採用する。不足している属性については、"ex:"のプレフィクスを用いて独自に定義する。PLODにおいて最も重要となる患者の移動情報の表現のためには、"schema:MoveAction"クラスに含まれる形で、"schema:startTime"、"schema:endTime"、"schema:fromLocation"、"schema:toLocation"のプロパティを用いる。データの真正性については、公衆衛生当局により発行された元のプレスリリースファイルのURLを明示することにより担保する。PLODモデルによって示した例をTurtle形式で表現したものを図3に示す。

4 評価

4.1 感染症患者発生に際したプレスリリースの分析

PLODによる患者情報公開の妥当性の検証のため、まず、検索エンジンを用いて、日本の感染症患者発生プレスリリースの収集を行った。その結果、2018～2020年の範囲で、厚生労働省と地方自治体が公開している麻疹189件、COVID-19を39件、合計228件のプレスリリースを収集することができた。これら収集したプレスリリースを対象として、患者情報がどのように記載されているのか、大まかな分類を試みた結果を表1に示す。

228件のプレスリリースのうち、71件に患者移動情報のための独立した節が設けられていた他、9件において、プレスリリースの本文中に移動に関する言及があった。また、さらに別の80件において、場所が特定されることを避けた一般名称を用いた形で、患者の移動情報が示されていた。16例では、不特定多数への接触リスクがない旨が記載されており、45件では、患者に関する詳細情報の記載が省かれていた。合計すると、228件中、224件に患者に関する記載が含まれ、そのうち80件に移動情報が含まれていた。これらのプレスリリースには、具体的な施設名の有無や、患者の移動を追えるだけの情報の有無等さまざまな点で違いがあり、発表情報の書式、内容ともに統一されていなかった。

なお、医療機関の場合は、施設名が公表されていないケースがあるが、医療機関側は当日の受診者と診察時間帯を把握しているため、感染リスクのあるものについては医療機関側から直接案内を出すことが出来る点に注意を要する。そのために、日本においては、プライバシー保護のために医療機関への訪問情報が意図的に省かれている。

```

1 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
2 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
3 @prefix schema: <https://schema.org/> .
4 @prefix dcterms: <http://purl.org/dc/terms/> .
5 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
6 @prefix gnjp: <http://geonames.jp/resource/> .
7 @prefix plod: <https://plod.info/property/> .
8
9 <https://plod.info/data/12202001311> a schema:Event ;
10   rdfs:label "12202001311" .
11
12 <https://plod.info/data/12202001311-R01> a schema:Report ;
13   rdfs:label "12202001311-R01" ;
14   schema:mainEntity <https://plod.info/data/12202001311> ;
15   plod:numberOfPatients "1"^^schema:Integer ;
16   schema:datePublished "2020-01-31"^^schema:DateTime ;
17   schema:publisher gnjp:Chiba ;
18   schema:url <https://www.pref.chiba.lg.jp/shippei/press/2019/ncov20200131.html>;
19   dcterms:isReferencedBy <https://www.pref.chiba.lg.jp/shippei/kansenshou/keihatu-index.html>.
20
21 <https://plod.info/data/12202001311-P01> a schema:Patient ;
22   rdfs:label "12202001311-P01" ;
23   schema:subjectOf <https://plod.info/data/12202001311> ;
24   schema:healthCondition <https://plod.info/entity/COVID-19> ;
25   plod:dateConfirmed "2020-01-31"^^schema:DateTime ;
26   foaf:age "20s" ;
27   schema:gender "Female" ;
28   schema:homeLocation gnjp:Chiba .
29
30 <https://plod.info/data/12202001311-P01-M01> a schema:MoveAction ;
31   rdfs:label "12202001311-P01-M01" ;
32   schema:agent <https://plod.info/data/12202001311-P01> ;
33   schema:startTime "2020-01-16"^^schema:DateTime ;
34   schema:endTime "2020-01-16"^^schema:DateTime ;
35   schema:fromLocation gnjp:Tokyo ;
36   schema:toLocation gnjp:Osaka ;
37   schema:instrument "Airplane"@ja .
38
39 <https://plod.info/data/12202001311-P01-M02> a schema:MoveAction ;
40   rdfs:label "12202001311-P01-M02" ;
41   schema:agent <https://plod.info/data/12202001311-P01> ;
42   schema:startTime "2020-01-22"^^schema:DateTime ;
43   schema:endTime "2020-01-22"^^schema:DateTime ;
44   schema:fromLocation gnjp:Osaka ;
45   schema:toLocation gnjp:Tokyo ;
46   schema:instrument "Bus"@ja .
47
48 <http://geonames.jp/resource/Tokyo> a schema:Place ;
49   rdfs:label "Tokyo" .
50
51 <http://geonames.jp/resource/Osaka> a schema:Place ;
52   rdfs:label "Osaka" .
53
54 <http://geonames.jp/resource/Chiba> a schema:Place ;
55   rdfs:label "Chiba" .
56
57 <https://plod.info/entity/COVID-19> a schema:InfectiousDisease ;
58   rdfs:label "COVID-19" ;
59   schema:name "2019-nCoV acute respiratory disease"@en ;
60   schema:infectiousAgent "2019-nCoV" ;
61   schema:code <http://purl.bioontology.org/ontology/ICD10/U07.1> .
62
63 <http://purl.bioontology.org/ontology/ICD10/U07.1> a schema:MedicalCode ;
64   schema:codeValue "U07.1" ;
65   schema:codingSystem "ICD-10" .

```

Figure 3: Example of PLOD (excerpt)

上記の分析は、検索エンジンを用いて集めた一部のプレスリリースに基づいている。プレスリリースの中には、各自治体のアーカイブを探さなければ得られないものもある。したがって、調査には網羅性が欠けており、分析には限界がある。しかし、筆者らが用意できたデータの範囲でみる限り、感染者の情報公開に際しては、地方自治体毎に発表数、発表情報の書式、内容ともに独自の形式で公開されており、標準化されたフォーマットに基づくものではない。英語圏のプレスリリースについても、検索エンジンを用いて12件のデータを確保し、確認したが、その範囲で見ると日本のプレスリリースと英語圏のプレスリリースに本質的な違いはなかった。したがって、公衆衛生当局により公開されるプレスリリースは標準化されておらず、感染者情報を統一的な書式により柔軟に公開しうるフレームワークの提案には十分な意義があるものと考えられる。

4.2 提案モデルの評価

次に、収集したプレスリリースのうち詳細な患者移動情報が含まれるものを用いて、PLODモデルの表現力の評価を試みた。今回は、2020/1/16~2020/2/11に公開された39件のプレスリリースを対象に、RDF (Turtle形式)を用いたPLOD化を実施した。プレスリリース自体の情報(メタデータ)、患者情報、患者の移動情報については、各エンティティに発行された永続的なIDをRDFの主語とした表現により、全件において問題なく表現することができた。ただし、詳細な情報が含まれないものについては、該当するRDFの目的語に該当する情報が空になることが確認された。39件のプレスリリースのうち、含まれる患者情報は23名であった。23名の患者のうち、日付と場所が明記された移動情報が含まれる患者は8名であった。

接触リスク計算のためには、移動元か移動先の情報があることが必須であり、より詳細な施設の情報があることが望ましい。移動情報がある8名の情報を、各移動のステップにより表現を分けると、21件の移動情報を抽出することができた。このうち、移動元か移動先が含まれるデータが10件であり、そのうち、都道府県レベルでの記載が8件、国外のみの記載のものが1件、緯度経度が特定できるような具体的な施設が記載されているものは1件であった。

また、移動情報が現在のPLODにて十分に表現できなかったものとして、以下の4パターンがあった。まず、①府外等、特定の地域以外であることが記載されているもの。次に、②バスガイド、運転手等、職業のみが記載されており、何らかの移動との関わりが示唆されるものの、症例の移動を追跡しうる粒度での情報が提供されていないもの。図4に、このパターンを例として記載する。さらに、③「小児と一緒に帰宅」等、テキスト情報からは直接位置情報を得られない場合がある他、④「過去の確定例と同様」、といった他のプレスリリースを参照する必要があるものが存在した。①、②のような場合には、表現から具体的な位置情報へと変換することが不可能である一方で、③、④のような例は、適切な処理により必要な情報が得られる可能性がある。たとえば、「帰宅先の自宅」という場合、文書を発行している自治体にその自宅が位置することが推測でき、他の文書を参照している場合、その文書を参照すれば位置情報の獲得が可能となる可能性がある。こうしたケースについては、RDFにより記述されたプレスリリースを対象とした推論処理によって明示されていない情報をも獲得しうるという点で、SPARQLに基づくアプローチに適したものと考えられる。

本評価の限界として、今回の評価は僅かなデータを対象とした分析に基づいており、ランダムサンプルができていないことが挙げられる。また、コーディング作業を1名で行っており、客観的な評価手法を用いることができていない。今後、客観的な評価手法の導入等、改善の余地がまだあるものの、提案手法により現在のプレスリリース中の情報のある程度表現しうるということが収集したサンプルより示された。

5 考察

本オープンデータ規格を日本において実用化する際には、いくつかのステップが考えられる。感染症対策は基本的に各自治体の責務であり、そのためにプレスリリース自体が各自治体からなされる。多くの場合は、自治体のウェブサイトにおいて、PDFかHTML形式で公開される。この状況において、携帯電話側のアプリケーションが、患者情報に関するオープンデータに効率的にアクセスするためには、大きく

```

1 <https://plod.info/data/29202001281-P01-M01> a schema:MoveAction .
2   schema:agent <https://plod.info/data/29202001281-P01> ;
3   schema:startTime "2020-01-08"^^schema:DateTime ;
4   schema:endTime "2020-01-11"^^schema:DateTime ;
5   schema:fromLocation "" ;
6   schema:toLocation "" ;
7   schema:instrument "Bus" .

```

Figure 4: Example of a failure case

わけて2通りの方法がある。1つ目は、個々のプレスリリースを厚生労働省等に集約したうえで、フォーマットを変換し、一般公開する中央集権的なアプローチとなる。もうひとつのアプローチは、各自治体が、RDFフォーマットと標準語彙を用いてそれぞれオープンデータを公開する分散的なアプローチである。後者のアプローチでは、携帯アプリケーション側が適切なデータを取得できるよう公的なRDFリポジトリ等の設置が必要となる。いずれにせよ、実現のためには行政機関をまたがったさまざまな調整が必要となるため、当面の方針として、まずは公開情報を第三者が手動で集積したうえでオープンデータ化し、実証実験目的で公開するというアプローチが現実的であろう。

提案手法の実利用に向けて次に必要となるのが、これらのオープンデータを活用して携帯ユーザーに感染リスクの通知を行う携帯電話側のアプリケーションの準備である。携帯端末には、位置情報の取得のためのAPI (Application Programming Interfaces) が備わっている。アプリケーションは、このAPIを定期的に利用し端末の位置情報を一定期間保存すれば、感染症患者の発生時に、互いの位置情報を利用して感染リスクの推計を行うことができる。アプリケーションが何らかのリスクを認めた場合、携帯電話の利用者に適切な警告を発することになる。

今回の仕様は、現時点では実アプリによる検証を経ていないドラフト水準のもので、実利用までにはいくつかのステップが存在する。本論でも言及したように、プレスリリースにはさまざまな種類の情報が記載されており、時折、現行のPLODにて表現することが困難なイレギュラーな情報が含まれることにある。そのために、今後、実際の活用を通じて、仕様自体をブラッシュアップしていく必要がある。一方で、プレスリリース側自体も、標準化に向けて改善されていく必要があるだろう。こうしたPLODとプレスリリース双方の発展に向けては、現行のPLOD仕様の表現力が限定されているとしても、まずは稼動する実験サービスを実現することが重要と考えられる。そうした実験サービスの利用を通じて、仕様とアプリケーションの双方が改善され、現段階ではクリアとなっていない様々な実用上のニーズへと応えていくことが可能となる。

ここで提案するアプリケーションの実装には、独立したアプリケーションとして実装するか、既に位置情報を取得している危機管理用のアプリケーションに組み込む形で実現するか等、いくつかのアプローチがある。提案手法はRDFに基づいているため、ユーザーや第三者は、自らのニーズに応じて様々な情報を付加することもできる。さまざまな開発者のアイデアを引き出すうえでは、ハッカソンを実施することも有益であろう。現在、感染症への感染リスクは世界的な関心事であるため、アプリケーションが容易に利用できるようになれば、広く携帯端末にインストールされていく可能性もあるだろう。

なお、パンデミックの状況下においては、アプリケーション利用者の増加に伴いオープンデータへのアクセスが急激に増加する可能性がある。そこで、サービス側も、危機管理用システムとして、高負荷時も安定運用できるようなスケーラビリティを備える必要がある。そもそも、最新のRDFストアでも、無数の携帯端末から来るSPARQLクエリの集中には耐えることは困難と考えられる。そこで、患者の移動情報オープンデータは、CDNs (Content Delivery Networks) 経由で配布するという手が考えられる。また、サーバの負荷を軽減するために、SPARQLクエリを簡素化するような工夫も可能かも知れない。SPARQLは、さまざまなデータの柔軟な検索を可能とするが、それは、サーバ側において複雑なクエリを処理する負担を負うことで実現されている。そのため、本稿で提案するような利用においては、引数をシンプルなものと保つことでより多くのクエリを処理しつつもSPARQLの柔軟性を維持する方向性が有用かも知れない。

6 結語

本稿では、感染症患者の状況に関するプレスリリース情報を、RDFを用いてオープンデータ化していくための枠組みを提案した。筆者らの知る限り、RDF技術とオープンデータを感染症リスクの軽減へと役立てる試みは知られていない。COVID-19により現在進行しているパンデミックによって技術に対する潜在的なニーズは多いものと考えられるため、今回提案した規格は標準として広く遣われる潜在力があるものと考えられる。

本研究の発展としては、今後、3つの段階が想定される。まずは、公開されているプレスリリースを収集し、研究として可能な範囲でオープンデータ化し、研究運用するRDFストアで公開する段階。これにより、技術的な完成度を高めると共に、感染症情報のオープンデータ化の有用性をさまざまな関係者に理解できる形とする。次のステップとして、発生していくプレスリリースを、公的なオープンデータとして公開していく枠組みへの移行がある。そのためには、厚生労働省や国立感染症研究所等の公的性質を有する機関において、全国のデータを集約してオープンデータとして公開する体制を実現する。そのうえで、最後のステップとして、各地方自治体自体が、自らの患者プレスリリースをオープンデータ、望ましくはLinked Open Data (LOD) として発行して頂く段階がある。

情報技術の公衆衛生への応用が進んでいるが、推計や予測といった分野に偏っており、感染症対応への活用は遅れている。また、医療におけるオープンデータ化とその活用事例は限られており、この分野の研究報告も数が限られている。我々は、今回の提案が、医療機関における診療負担増の警告手段等、公衆衛生分野における知識処理とオープンデータ化の手本となっていくような展開を期待している。

References

- [1] C. Yang, J. Yang, X. Luo, and P. Gong, "Use of mobile phones in an emergency reporting system for infectious disease surveillance after the sichuan earthquake in china," *Bulletin of the World Health Organization*, vol. 87, pp. 619–623, 2009.
- [2] K. Kim, T. Tandi, J. W. Choi, J. Moon, and M. Kim, "Middle east respiratory syndrome coronavirus (mers-cov) outbreak in south korea, 2015: epidemiology, characteristics and public health implications," *Journal of Hospital Infection*, vol. 95, no. 2, pp. 207–213, 2017.
- [3] K. H. Jones, H. Daniels, S. Heys, and D. V. Ford, "Toward an ethically founded framework for the use of mobile phone call detail records in health research," *JMIR mHealth and uHealth*, vol. 7, no. 3, p. e11969, 2019.
- [4] B. Ouedraogo, J. Gaudart, and J.-C. Dufour, "How does the cellular phone help in epidemiological surveillance? a review of the scientific literature," *Informatics for Health and Social Care*, vol. 44, no. 1, pp. 12–30, 2019.
- [5] S. Lai, A. Farnham, N. W. Ruktanonchai, and A. J. Tatem, "Measuring mobility, disease connectivity and individual risk: a review of using mobile phone data and mhealth for travel medicine," *Journal of travel medicine*, vol. 26, no. 3, p. taz019, 2019.
- [6] K. H. Jones, H. Daniels, S. Heys, and D. V. Ford, "Challenges and potential opportunities of mobile phone call detail records in health research," *JMIR mHealth and uHealth*, vol. 6, no. 7, p. e161, 2018.
- [7] E. Jahani, P. Sundsøy, J. Bjelland, L. Bengtsson, Y.-A. de Montjoye *et al.*, "Improving official statistics in emerging markets using machine learning and mobile phone data," *EPJ Data Science*, vol. 6, no. 1, p. 3, 2017.
- [8] J. Chirombo, P. Diggle, D. Terlouw, and J. Read, "A review of models of human mobility for predicting infectious disease spread," *Modelling spatial processes of infectious diseases*, p. 60, 2018.
- [9] C. Panigutti, M. Tizzoni, P. Bajardi, Z. Smoreda, and V. Colizza, "Assessing the use of mobile phone data to describe recurrent mobility patterns in spatial epidemic models," *Royal Society open science*, vol. 4, no. 5, p. 160950, 2017.

- [10] K. Sallah, R. Giorgi, L. Bengtsson, X. Lu, E. Wetter, P. Adrien, S. Rebaudet, R. Piarroux, and J. Gaudart, “Mathematical models for predicting human mobility in the context of infectious disease spread: introducing the impedance model,” *International journal of health geographics*, vol. 16, no. 1, p. 42, 2017.
- [11] A. Wesolowski, C. O. Buckee, K. Engø-Monsen, and C. J. E. Metcalf, “Connecting mobility to infectious diseases: the promise and limits of mobile phone data,” *The Journal of infectious diseases*, vol. 214, no. suppl_4, pp. S414–S420, 2016.
- [12] L. Bengtsson, X. Lu, A. Thorson, R. Garfield, and J. Von Schreeb, “Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: a post-earthquake geospatial study in haiti,” *PLoS medicine*, vol. 8, no. 8, 2011.
- [13] F. Finger, T. Genolet, L. Mari, G. C. de Magny, N. M. Manga, A. Rinaldo, and E. Bertuzzo, “Mobile phone data highlights the role of mass gatherings in the spreading of cholera outbreaks,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 23, pp. 6421–6426, 2016.
- [14] J. Cinnamon, S. K. Jones, and W. N. Adger, “Evidence and future potential of mobile phone data for disease disaster management,” *Geoforum*, vol. 75, pp. 253–264, 2016.
- [15] L. Bengtsson, J. Gaudart, X. Lu, S. Moore, E. Wetter, K. Sallah, S. Rebaudet, and R. Piarroux, “Using mobile phone data to predict the spatial spread of cholera,” *Scientific reports*, vol. 5, p. 8923, 2015.
- [16] M.-P. Gagnon, P. Ngangue, J. Payne-Gagnon, and M. Desmartis, “m-health adoption by healthcare professionals: a systematic review,” *Journal of the American Medical Informatics Association*, vol. 23, no. 1, pp. 212–220, 2016.
- [17] L. Lin, “China Marshals Its Surveillance Powers Against Coronavirus,” <https://www.wsj.com/articles/china-marshals-the-power-of-its-surveillance-state-in-fight-against-coronavirus-11580831633>.
- [18] E.-Y. Jeong, “South Korea Tracks Virus Patients’ Travels—and Publishes Them Online,” <https://jp.wsj.com/articles/SB12291155354026644516304586207690851702666>.
- [19] T. Okumura, “Tracing infectious agents with mobile location information: A simple and effective countermeasure against epidemic risks,” in *2019 IEEE Global Humanitarian Technology Conference (GHTC 2019)*. IEEE, 2019.
- [20] C. Torniai, S. Battle, and S. Cayzer, “Sharing, discovering and browsing photo collections through rdf geo-metadata,” in *SWAP 2006 - Semantic Web Applications and Perspectives, Proceedings of the 3rd Italian Semantic Web Workshop*, ser. CEUR Workshop Proceedings, vol. 201. CEUR-WS.org, 2006. [Online]. Available: <http://ceur-ws.org/Vol-201/07.pdf>
- [21] R. Battle and D. Kolas, “Geosparql: enabling a geospatial semantic web,” *Semantic Web Journal*, vol. 3, no. 4, pp. 355–370, 2011.
- [22] D. Beckett and B. McBride, “Rdf/xml syntax specification (revised),” *W3C recommendation*, vol. 10, no. 2.3, 2004.
- [23] E. Prud’hommeaux and A. Seaborne, “SPARQL Query Language for RDF,” W3C Recommendation, 2008. [Online]. Available: <http://www.w3.org/TR/rdf-sparql-query/>