Vol.2009-EIP-45 No.2
2009/9/12

プローブ情報システムへの匿名性評価手法導入のための一考察

和泉順子† 佐藤雅明†† 砂原秀樹†, †††

インターネット等の共通情報通信基盤上に構築されるプローブ情報システムは国内外からの高いイノベーションが期待される一方，車両追跡や行動履歴の推測などの脅威が存在する．現在，各種プローブ情報システムの脅威分析や個人情報保護に関するガイドラインは国内外で検討されているが，プローブデータ自体の信頼性やシステムセキュリティの検証に関する議論は少ない．そこで，汎用性，社会性の高いプローブ情報システムが扱う情報の信頼性やシステムセキュリティを評価・検証手法を考察する．

A Study on the Introduction of Evaluation Criteria for Anonymity on Probe Vehicle Information Systems

Michiko IZUMI† and Masaaki SATO††
and Hideki SUNAHARA†, †††

Probe vehicle information systems which gather various sensor data, called probe data, by some sensors or equipments on vehicles can innovate new methods to build not only the fine-grained social information services, but also some valuable business around the ITS (Intelligent Transport Systems) area. Probe data for wide area communications were standardized internationally, however, it's noted that there are diversification of threats to privacy because that the collected probe data would include the spatiotemporal data when the vehicle obtained the information. These raw data can create a personal routing or action record, thus appropriate techniques and those threat assessments, which can evaluate through common criteria, are much in demand on this area.

In this paper, we discuss the provision for privacy issue, especially anonymity, on the probe vehicle information systems. There exists a tradeoff between data quality (or its utility) and the degree of privacy, however, we should carry the weight of versatility for some existing services and scalability for the future.

1. はじめに

Mobile IPv6 等の次世代インターネット技術を含めた多用な通信手段を用いて構築されるプローブ情報システム(PVS: Probe Vehicle Systems)，FCD(Floating Car Data)と呼ばれるサービスは，車両に搭載されたセンサから走行状態や位置・時刻・周辺環境を取得し，きめ細やかな交通情報等の統計情報を生成することが可能である．HONDA のインターナビ・フローティングカーシステム[1]や TOYOTA のプローブコミュニケーション交通情報[2]，BMW の XFCD(Extended Floating Car Data System)[3]のように多様なサービスが既に実用化され，国内では VICS (Vehicle Information and Communication System)情報に連動させるなど有効活用されているものも多い．しかし，これらの情報は現在ではサービスを提供する事業者内で閉じて管理運営しており，通信手段やデータ管理方法等その形態は様々である．

しかし，このような情報を共通の情報通信基盤で流通させることができれば，渋滞情報や気象情報だけでなく，車両運行予測や旅行者サポート等の道路・交通の管理や情報提供サービス提供が可能になるため，イノベーションが期待される重要な技術領域として認識されており，車両から提供されるデータ形式については ISO において IS22837[4]として国際標準化が成されている．この標準で定義され，実際に利活用されるプローブデータ自体は統計情報であり個人情報の収集は必要とされておらず，国際標準データ群には個人情報は含まれていない．しかし，通信 ID やプローブデータの性質上，一定時間の観測や複数情報の統合により個人情報やプライバシに関わる情報も類推可能となる等の脅威も存在することが知られており[5][6]，関連ガイドライン策定の場でも議論が続けられている．

つまり，プローブ情報システムから生成される各種情報サービスは有益であると認識され，国内外で注目される一方，車両からのプローブデータ提供者については自動車とその所有者の関連性が深いため，広くサービスを開発・展開するにはシステム自体の安全性や個人情報保護への対策が不可欠であり，対策が急がれる．車両からプローブデータを提供する利用者は，個人データの秘匿性や事業者に対する信頼，つまり技術的な情報システムそのものの安全性やデータ利用の信頼性と，運用的な観点からの事業者の対応等を客観的に検証できる状態でなければ，プローブ情報を発信するこ

† 奈良先端科学技術大学院大学 情報科学研究科
Graduate School of Information Science, Nara Institute of Science and Technology, Japan
†† 慶應義塾大学大学院 政策・メディア研究科
Graduate School of Media and Governance, Keio University
††† 慶應義塾大学大学院 メディアデザイン研究科
Graduate School of Media Design, Keio University

とに不安を感じ，情報提供を躊躇してしまう．

このような状況を打開するためには，閉じたサービスではなく共通の情報通信基盤上を流通するプローブデータを活用することの有効性を示すと同時に，提供されるデータがどのような属性を持ち，匿名性などの安全面はどう保たれるのか，ということを客観的に示すことができる共通指標の導入が必要となる．また共通指標の導入は，他事業者，または他業種間でのサービス協調を図る際の判断材料にも有効であり，対外的に明確にサービスを説明可能となるため，サービス提供者にとっても有益なものとなる．

本論文では，共通情報通信基盤上で流通するプローブデータを用いて多様なサービスを発現させるために，車両からプローブデータを提供する利用者に，どのような情報がどのように提示されるべきか，について考察する．第２章ではプローブ情報システムに関連する技術および標準化動向を概説し，第３章で既存セキュリティ評価基準および情報量（エントロピー）などの概念を用いた関連指標について述べる．それらを踏まえて第４章でプローブ情報システムに必要なセキュリティの内，特にデータの匿名性確保について導入すべき評価指標を考察し，最後に第５章で議論をまとめる．

2. プローブ情報システム

プローブ情報システムは，情報通信基盤を用いて自動車の持つセンサ情報，あるいは周辺状況を収集し，その情報に統計処理を施すことで有益な情報サービスを生成する，センサ情報共有・提供のためのシステムである．したがって，このシステムで扱うプローブ情報とは，車両を通じて収集されるデータとなる．図 1 にプローブ情報システムの概略を示す．車両のセンサや搭載された機器から収集されるデータは，携帯電話網や無線 LAN などの情報通信基盤を用いてプローブ情報センタに送信される．プローブ情報センタで位置・時刻・路面状況のような周辺情報等をプローブデータとして正規化して収集・加工することにより，渋滞情報や気象情報，運行予測や旅行者サポート等の，道路・交通の管理や情報提供サービス提供が可能になる．

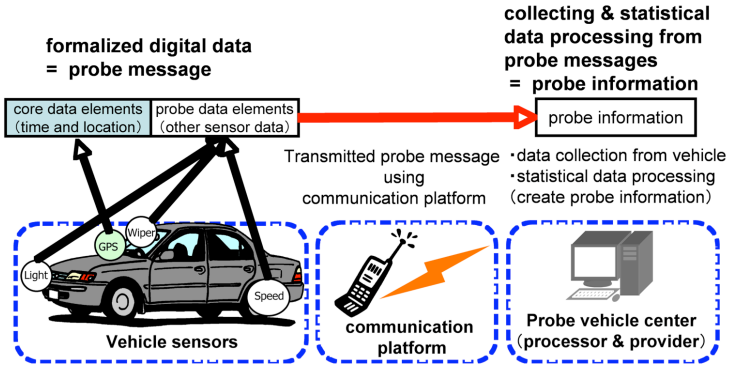

図 1　プローブ情報システム

現在，目的や用途毎に国内外で研究開発や実証実験，実事業化が行われているプローブ情報システムは数多く存在しており，個々のシステムの安定動作だけでなく複数サービスの協調や情報共有を検討する上では，共通の情報流通基盤が必要となる．この基盤構築には，

- 情報提供者が安心して情報を提供できる
- 情報収集者が適切かつ容易に情報を共有できる

という二つの条件が必要となる．この二つを満たすための客観的な指標，特にプローブデータの匿名性評価手法導入の検討が本論文の目的である．本節では，プローブ情報システムの概説として関係する国際標準と，プローブ情報システムにおける脅威について述べる．また，2008 年 11 月に行われた ITS 世界会議で得られた知見として ITS(Intelligent Transport Systems : 高度道路交通システム)分野での情報システムのセキュリティ技術への関心の高さを示すと同時に，情報提供者（以下，利用者）の不安とプローブ情報システムに対する事業者の関心についてのアンケート調査結果から導出される傾向を掴む．

2.1 プローブ情報システムの動向

プローブ情報システムは，車両に搭載されたセンサから走行状態や周辺環境を取得し，交通情報等の統計情報を生成することが目的であるため，個人情報の収集は必要とされていない．情報流通基盤上で取り扱うデータ形式やリファレンスモデルはプローブ情報の国際標準データ群国際的に標準化[4]されており，このデータ群には個人情報は含まれていない．

しかし，プローブ情報に付加された時刻情報や物理的な位置情報，およびその履歴

等を用いることで車両を特定または推定することが可能であり，かつ車両とその所有者との結びつきが強いことから，車両追跡や行動履歴の検索の可能性などのプライバシ侵害につながる脅威が存在する．通信に用いる ID の関連性やサービスを享受する際の認証処理に用いる属性情報等から考えられる脅威や対策については，文献 [5],[6] で整理され議論されている．さらに，個人情報が公開されプライバシが侵害された場合の損害が甚大になる場合が多いため，このような情報の取り扱いには特に慎重な運用・配慮が求められている．しかし，現状としては技術的な対応として極めて高度なセキュリティ機能をサービスに盛り込む場合，利用者だけでなく情報管理者にとっての利便性や扱う情報そのものの精度を損なうことも多いため，個人情報そのものやプライバシの侵害に関するリスクを極力低減するため，大抵の場合は技術的な対策よりも機能やサービスを限定する等の運用面での対策を採用し，情報共有やサービスが独立または限定されていることが多い．また，その安全性を客観的に検証，追認することが困難であり，実現可能な具体的対策とその評価指標については未だ議論がまとまっていない．

このように情報システムにおける個人情報の取り扱いは国際的にも関心を集めており，プローブ情報に関する脅威分析や個人情報保護に関するガイドライン作りが ISO/TC204/WG16[7]内のサブグループである SWG16.3 で検討されている．

この SWG ではプローブサービス事業者が対応すべき事項を，OECD (Organization for Economic Co-operation and Development：経済協力開発機構)が，「収集制限の原則」や「目的明確化の原則」,「安全保護の原則」などの 8 原則で規定している「プライバシ保護と個人データの国際流通に関するガイドラインに関する理事会勧告[8]」を参考にした基本原則として各国の状況を鑑みながら議論・検討が重ねられている．この SWG での議論の結果は国際標準案としてまとめられ，平成 16 年 10 月に ISO へ PWI 24100 として提案された．その後，この標準案は規定の標準化段階を経て，DIS 投票が終わり，現在は FDIS としての登録を承認された[9]状態となっている（平成 21 年 8 月 1 日現在）.

2.2 ITS 分野における情報セキュリティへの関心

2008 年 11 月に米国で開催された第 15 回 ITS 世界会議（15th World Congress on ITS）のテーマは"ITS Connections: Saving Time, Saving Lives"であり，Executive session, Scientific session, Special Session, Technical session など，規模もテーマも様々な 300 以上のセッションが行われた. その中でも,「Data Security and Privacy」「Communication Technology for ITS」「Social Implications for ITS Data Collection」等のセッションではプローブ情報の活用には大きな関心が寄せられており，日本だけでなく，欧米やアジアからも多くのプローブ関連研究者や事業者が研究発表や展示を行っていた. また, これらのセッションではプローブ情報システムへの関心と同様に，システムのセキュリティや扱うデータに対するプライバシ保護についての関心も非常に高く，深い議論も随所でなされていた.

これらのセッションの共通点として，プローブ情報システムへの関心と同様に，システムのセキュリティや扱うデータに対するプライバシ保護についての関心も非常に高かった．データやシステムの安全性や匿名性の維持は，データが本来持つ精度を低減させたり，サービスの利用者または運用管理者に対するシステムの利便性・操作性を損なったりすることも多く，これらは往々にしてトレードオフの関係として認識されている[10],[11]．中でも位置情報を扱う既存サービス・システムに対するプライバシ保護を anonymity（以下，匿名性）の保証として問題を扱い，この対策と評価に関する発表には活発な質疑応答により議論が展開されていた．本論文でも主に，匿名性に関する評価手法および指標について議論することとする.

2.3 利用者の不安，事業者の関心

事業者のプローブ情報への関心については，経済産業省が主催する基準認証事業の活動の一環であるワークショップで実施されたアンケート[12]からも明らかである．さらにこのアンケートでは，利用者のデータ提供に関する不安傾向についても概略を掴む事ができる.

このワークショップは 2006 年 12 月 19 日に，基準認証事業「プローブ情報システムにおける個人情報保護に関する標準化」の活動内容と策定したガイドラインの紹介等を目的として開催された．ここでは ITS 関係者，特にプローブ情報サービスの利用に関心がある方に参加を呼びかけ，ITS に関連する個人情報保護について，事業者として，または利用者として，どのような関心・問題を抱えているかを調査するためアンケートを実施しており，ワークショップに出席した一般参加者 64 人の内，回収率 88%となる 56 人から回答を得ている．回答の結果，主に以下の様な傾向が掴めた.

- 求められれば提供するプローブデータ
  - 「ワイパー」,「ヘッドライト」,「ABS」等
  - 逆に「カメラ（画像）」や「経路情報」は求められても提供に同意できないという傾向が強い
- 上記設問の提供に同意できない理由
  - 最多は「プライバシ情報だから」「どのように利用されるかわからないから」
  - 「はっきりした理由はない」という選択肢を選んだ人はいない
- 提供同意に必要な条件
  - 最多は「不愉快な思いをしないことが保証される」
  - 他にも「社会的な効用がはっきりしている」「事業者が信頼できる」等

も多かった

- プローブデータの利用に関心があるか
  - 「はい」93%
  - 「いいえ」2%,「どちらともいえない」4%
- 利用したいプローブデータの属性
  - 最多は「個人が特定できなくてもよい」の 75%
  - 他に「個人が特定できなくてもよいが,年齢と性別は特定したい」20%,「個人が特定できる」11%

つまり，利用者としては，データを提供することで不愉快な思いをしないことを前提に，提供データの使われ方や効果，事業者の信頼度や対応等を客観的に提示することができれば，データの提供同意に大きくつながると予測できる．また，事業者としては，マーケティング等に有効であるため一部「個人が特定できる」または「年齢と性別が特定できる」データの需要があるが，多くは個人が特定できないプローブ情報でも十分利用に関心がある，ということがわかる．

国際標準化の動向から，目的や用途毎に国内外で研究開発や実証実験，実事業化が行われているプローブ情報システムの協調や情報共有を検討する上で必要な共通の情報流通基盤の需要と期待の高さが読みとれる.しかし,その期待を実現するためには，共通の情報通信基盤を流通するデータやシステムの安全性や匿名性の維持が不可欠である．これについては位置情報を用いたインターネット技術，特に移動体通信関連技術において，データの匿名性やシステムの安全性が，データが本来持つ精度の低下やシステムの利便性・操作性の喪失につながり，トレードオフの関係として認識されていることが分かったが，それでもなお，研究開発としても事業化としても匿名性を確保したプローブデータに関する関心は高いことがわかった．

次節移行で，プローブ情報システムを展開する上で不可欠なセキュリティ評価基準の動向と，情報科学分野において匿名性を確保する技術について議論する．

3. 既存の評価基準および評価手法

プローブ情報システムを広く開発・展開するにはシステム自体の安全性やプローブデータの匿名性確保よる個人情報保護への対策が不可欠である．そこで，本節ではプローブ情報システムのプライバシ対策・セキュリティに対する客観的な評価指標の必要性を議論するために，既存の枠組みとして IT セキュリティ評価基準に関係する国際規格等を調査し，問題点および評価指標整理の参考とする．

まず，参考とすべき評価基準としては，ISO/IEC15408 情報技術セキュリティ評価のためのコモンクライテリア(Common Criteria for Information Technology Security Evaluation)[13],[14],[15]，およびそれに関連する情報技術セキュリティ評価のための共通方法論が挙げられる.これらは非常に有益で汎用的な国際規定であり，サービスを提供する企業にとっても，自サービスの守備範囲，想定利用方法等が国際標準に準拠した形で明確に記述できれば，利用者に明確で客観的な説明ができるだけでなく，他サービスと協調する時の判断基準にもなり得る．しかし，実際にこの標準に則って個々のプローブ情報システムのセキュリティ・プライバシ機能を評価・認証をするには準備したりプローブ情報システムのプロテクションプロファイル（PP）を策定・提案したりするには莫大な時間と費用が必要となる．したがって，実際の認証申請準備をするのではなく，この規定を参考にするに留めるが，この規定の整理を参考にすることにより「プローブ情報に関する脅威分析や個人情報保護に関するガイドライン[9]」について，より汎用的で客観的な整理が可能となると考える．

また，2.2 項で述べた通り，本論文ではセキュリティ・プライバシ機能に関するものの中で特に匿名性について議論する．利用者が，サービスやアプリケーションに関する個人情報の取り扱いを客観的に比較・検証できる手法や指標として，既存技術の中でプローブ情報システムに適応可能であるものを検討する．

3.1 情報セキュリティに関する評価基準

情報技術製品のプライバシ保護を議論する中で重要な概念となる「anonymity（匿名性）」は，ISO15408-2:2008（情報技術セキュリティ評価のためのコモンクライテリア Part2 セキュリティ機能要件）[15]で定義される機能要件，具体的にはプライバシに関する機能要件クラス(FPR)のファミリの 1 つでもある．クラスには「クラス名」「クラスの概説」一つ以上の「機能ファミリ」が含まれており，機能ファミリには「ファミリ名」「ファミリの振る舞い」「コンポーネントのレベル付け」「管理」「監査」コンポーネント」が含まれる．機能ファミリ内のコンポーネント間の関係は階層関係になることもあり，そのレベル付けもここで定義される．ここに定義されているプライバシクラス(FPR)では，プライバシ機能要件として利用者の識別情報を他の利用者による開示，及び悪用から保護することを設けており，コンポーネント構成は図 2 ように「anonymity（匿名性）：FPR_ANO」「pseudonymity（偽名性）：FPR_PSE」「unlinkability（リンク不能性）:FPR_UNL」「unobservability（観測不能性）:FPR_UNO」の 4 つの機能ファミリにカテゴライズされ，それぞれが機能コンポーネントを含んでいる形になっている．

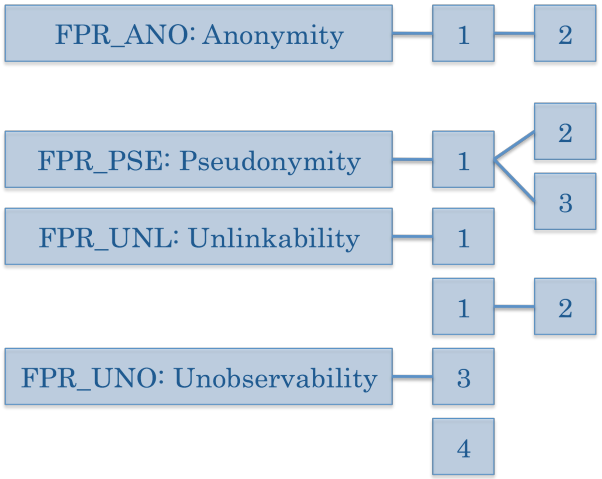

図 2　プライバシクラス(FPR)のコンポーネント構造

本研究では，プローブ情報システムのセキュリティ・プライバシ評価基準について，この構造に即した形で整理が必要になる．ここでいう Anonymity（匿名性）とは，あるサブジェクトまたは操作に結びつけられた利用者の識別情報を，他の利用者やサブジェクトが判別できないこと（Anonymity, requires that other users or subjects are unable to determine the identity of a user bound to a subject o r operation）であり，ある一つの情報が，誰によるものかが不明である状態，と定義されている．また，2 つあるコンポーネントの内，FRP_ANO.1 匿名性は，「あるサブジェクトまたは操作に結びつけられた利用者の識別情報を，他の利用者やサブジェクトが判別できないことを要求する」となっており，FRP_ANO.2 では「情報を請求しない匿名性は，TSF が利用者識別情報を要求しないことを保証することによって，FRP_ANO.1 匿名性の要件を強化する」と定義されている．TSF とは TOE（評価対象）セキュリティ機能のことである．したがって，匿名性の議論については「あるサブジェクトまたは操作に結びつけられた利用者の識別情報を，他の利用者やサブジェクトが判別できないこと」を評価軸の一つとして議論する．

また，情報技術セキュリティ評価のためのコモンクライテリアのプライバシクラス構造に従うには，匿名性の他に，Pseudonymity（偽名性），Unlinkability（リンク不能性），および Unobservabiltiy（観察不能性）に対する議論が必要となる．偽名性は，利用者が資源やサービスを利用するにあたり，利用者識別情報を開示する必要はないが，その利用に関する説明責任を有することを保証することについて，リンク不能性は，利用者は複数の資源やサービスを利用できるが，他の利用者はこれらの資源やサービスの利用を関連付けることができないことを保証することについて，観察不能性は，利用者が，資源やサービスを利用にあたり，他の利用者，特に第三者が，その利用を知ることができないことを保証することについてそれぞれ定義し，その方法について議論と整理をする必要があるが，これらについては本論文の範囲外とする．

3.2 匿名性に関する評価手法

情報科学分野において，プライバシを議論する上で重要な概念となる匿名性の概念に基づいてプライバシ保護を試みる専攻研究が，匿名通信および位置情報サービスの分野で多数行われている．匿名通信とは，ネットワーク上を流れる情報だけでは通信の送受信者が判別できない様な通信の形態である．匿名性を確保するための手段・技術の一つであり，1981 年に Chaum が発表した Mix-net [16]や，2003 年 IEEE Pervasive computing 誌の Frank Stajano 氏の記事[17]にある Mix zone などに代表される．これは受信者と送信者との間に MIX と呼ばれるサーバあるいは zone を設け，これらの中継点がここを通過する複数のメッセージをかきまぜることでメッセージの流入経路等を隠蔽する手法である．これは，送信メッセージと受信メッセージの相関を隠蔽するために，ランダムな遅延に基づく時間的な曖昧さ等を発生させ，その相関を小さくする手法であり，メッセージ間の非結合性つまりはリンク不能性（Unlinkability）を高めることによりメッセージの匿名性を実現するとも云える．

また，他にも匿名性の概念に基づいたプライバシ保護技術としては，匿名認証（Anonymity authentication / Anonymity Credential），k−匿名性（k-anonymity ）といった技術も利用されつつある．文献[10]やでは，プローブ車両の匿名性を確保するために k−匿名性の技術を採用しており，その実用性を実験により評価している．匿名認証技術は，ゼロ知識証明やグループ署名，ブラインド署名などの技術エッセンスを組み合わせた技術であり，論文[18]や，無線モバイルアドホックネットワークにおける安全な通信プロトコルの提案をしている技術論文[19]において採用実績がある．これらの匿名性の検証については，実験または数学的な理論評価がなされている．

このように，プライバシを確保するために匿名性を確立する技術は複数存在する．これらの技術はそれぞれが対象とするサービスで匿名性を保証するものであり，技術自体が独立した概念で確立されているため，このままの状態では匿名性を客観的に共通の指標で検証することはできない．また，[10]においては，k-匿名性に基づく既存アルゴリズムが location traces（位置追跡/捕捉）を大幅に修正することもあるため，データ精度に関する要件が満たされなくなってしまうおそれがある，とも記述されている．つまり，同じ技術を用いる場合でも，サービス全体としてはデータの疎密（バラツキ）や時間軸，生起確率等，複数変数から構成されているため，システムを設計する上でそれらの関連づけや取り扱い次第ではプライバシ保護以外の要件が満たされ

なくなる. これらの議論より, 匿名性を保証するための技術を単体で使うのではなく, 複数の既存技術や他の統計的または数学的手法を組み合わせて, 匿名性を保証するシステムを設計・構築することも考えられる.

プローブ情報システムは, 通信形態や利用する機器が多種多様であり, 求められるセキュリティレベルもまちまちであるため, システム設計も, 匿名性確保のための仕組みや組み合わせも多岐にわたる. このような場合においても匿名性を客観的に, かつ定量的に評価できる指標を検討する必要がある.

3.3 匿名性の評価指標

ここでは, システムやデータのセキュリティ, あるいはプライバシ保護に対する一般的な評価指標について議論する. 文献[17]の著者である Frank Stajano 氏は, 英国ケンブリッジ大学コンピュータラボラトリの Tenured Senior Lecturer であり, モバイル, ユビキタスネットワークのセキュリティに造形の深い研究者である. 多くの技術論文がこの文献を参照しており, 当該分野における影響は大きいと考えられる. この論文中には, 匿名性に関する評価指標として, 以下の二つが導入されている.

- Anonymity Set
- 情報理論のエントロピー（情報量）

(1) Anonymity Set

Anonymity Set は, 文献[17]では"The anonymity set's size is a first measure of the level of location privacy available in the mix zone at the time", つまり, Anonymity set のサイズは mix zone における利用可能なロケーションプライバシのレベルに対する最初の評価尺度であると定義されており, 匿名性やリンク不能性などの用語集をまとめている文献[20]では, “Anonymity of a subject means that the subject is not identifiable within a set of subjects, the anonymity set. The anonymity set is the set of all possible subjects”と定義されている. Anonymity Set は特定のメッセージに対してそのメッセージを送信する可能性のある送信者の集合, つまり, 個人またはエンティティを特定できない集合（mix zone 等）におけるそのエンティティの数, として定義されており, この数が大きければ大きい程匿名性は高い. 例えば, Anonymity Set が 20 の mix zone よりも Anonymity Set が 1,000 の mix zone の方が匿名性が高いのは直感的に自明である. 特定のメッセージを追跡しようとする攻撃者には, メッセージを送ることができる送信者が 1 人しかいない場合, 容易にそのメッセージの送信者を特定することができるが, メッセージを送ることができる送信者が多ければ多いほど, その特定が困難になる. つまり, 特定のメッセージを送信可能である送信者の数が多ければ多いほど, そのメッセージの匿名性が高いと云える.

位置情報プライバシに関する重要な概念である匿名性の評価については, Anonymity Set の概念が広く浸透しており, 必要に応じて適宜拡張や応用がなされているが, Anonymity Set で評価できるのは, その構成要素の生起確率について蓋然性が等しい場合だけである. 実際には匿名性の概念は, mix zone を通過する際の侵入または退出の位置情報（つまり, 何処から来て何処に進むか）に強く関係するため, 評価指標として Anonymity Set だけでは不十分であると云える.

(2) 情報理論のエントロピー

前項の Anonymity Set の欠点を受けて, 文献[17]では, 匿名性の評価尺度としてシャノンの情報量（以下, エントロピー）を導入している. これは, Anonymity Set の概念に確率分布の要素を定義することで利用できる匿名性の評価指標である. ここでは, 生起確率が等しくなければエントロピーは小さくなる性質を利用している.

事象 E が起こる確率を P(E)とするとき, 事象 E が起こったことを知らされたとき受け取る（選択）情報量 I(E) は, 以下の様に定義される.

$$I(E) = \log \frac{1}{P(E)} = -\log P(E)$$

この定義からも, エントロピー I(E)は明らかに, 生起確率が低い事象の情報量ほど値が大きくなることが分かる. エントロピーが増大するということは不確実性が高まるということを意味しており, この値が大きいほど匿名性が高い（信頼性が高い）と云える. つまり, 生起確率が低い方が攻撃者からも追跡・特定が困難になるということを示しているため, 直感的な認識とも矛盾を生じない.

なお, 文献[21]によると, シャノン情報量は平均的な挙動の場合に用いる指標であり, 最悪の場合の指標としてはシャノン情報量は適さない（misleading）と示されている. この場合は最小エントロピーを用いることが提唱されており, また global anonymity の指標としては上界は最大エントロピーであるため, この両界を埋めるものとして Renyi エントロピーの利用が提案されている.

4. プローブ情報システムに必要な匿名性の確保と評価指標

ここでは, 前節までに述べた匿名性の評価手法や指標をプローブ情報システムに適用するための議論を行う. まず, プローブ情報システムで扱われるデータ属性として必要な匿名性について 2 章の議論を基に検討する. 検討により導出される要件を 3 章で述べた既存評価基準や匿名性評価指標でどのようにサポートするべきか, また, 今後どのような検討が必要かについて議論する.

4.1 プローブ情報システムに必要な匿名性の性質

2 章で述べた通り, プローブ情報システムは国内外で広く研究開発されており, プ

ローブデータの形式や個人情報保護に関するガイドラインについては標準化が進んでいることからも期待と需要の高い分野である．しかし，これらは各国の法令または通信事情やサービスを開発･提供する事業者の目的や規模,用途等が多岐にわたるため，扱うプローブ情報の属性として「匿名性」の確保とその検証可能性が求められているとはいえ，全てを同じレベルで評価することは現実的ではない．たとえば，国内幹線道路の交通流量制限や信号制御などの道路交通情報管理に関わるプローブデータとしては，利用者は高い匿名性を維持して必要最小限のデータ以外は送信を控えたいと考えるかもしれないが，特定の自治体内を走行するバスの運行管理等では，利用者（この場合はバス運転者）は個人の車両ではなく業務車両であるため，そこから発信されるプローブデータには匿名性があってもなくても良いと考えるかもしれない．

したがって，共通の情報通信基盤上を流通するプローブデータがどのような匿名性という属性を如何に実現するかを客観的に示す共通指標には，必要な匿名性強度と，その達成度合い(degree of anonymity)という 2 次元の指標が必要であると考える．

達成度合いは，必要となる匿名性強度とシステムの規模により採用する匿名化技術（匿名認証，ランダム ID 等）を選択することによって整理可能であると考える．このそれぞれの技術について，匿名性の達成度合いとしてエントロピー等の概念を用いて定量的な数値が提示できれば，共通情報通信基盤上でのプローブデータの匿名性を検証するという目的達成に大きく近づく．

匿名性強度についてはそれぞれのシステムで求められるものが違うため，整理・検証のための枠組みを検討するに留めざるを得ない．この枠組みは，既存システムが適応可能なカテゴリ分けでありながら，将来性を見据えた汎用性が必要とされるため，今後プローブ情報システム関連の研究開発と標準化に携わる関係者間で広く議論が必要となる．

4.2 匿名性に関する評価手法

3 章で，参考にすべき情報セキュリティ評価基準と情報科学分野での一般的な匿名性評価の概念について述べた．匿名性強度に関する枠組みは互いに独立しており，それぞれにおいて達成度合いを FRP_ANO.1 匿名性の「あるサブジェクトまたは操作に結びつけられた利用者の識別情報を,他の利用者やサブジェクトが判別できないこと」を評価軸として採用可能技術ごとに整理する必要がある．匿名性の達成度合いについては，Anonymity Set やエントロピーの概念を用いることで客観的な指標を示す事も可能であると考える．

しかし，文献[20]では，システム全体の匿名性を高く保てたとしても，その中の一つの生起確率だけが極めて高い場合，匿名性が強いとは云えないことと指摘し，問題提起している．この文献では global / individual という言葉を用いているが，文献[21]では同様の議論に global/ local の anonymity という言葉で整理をしている．つまり，収集した全体のプローブデータの匿名性を議論するだけでなく，その中での生起確率やセンサ値が突出している(周囲のデータと比較して明らかに違う値になる)場合は，このデータを除外する等の適切な処理が必要となる．

生起確率やセンサ値が異常値を出すのは，センサ異常や悪意のある第三者による攻撃，あるいは過疎地における情報提供（つまり個人が特定しやすい状態）などの状況が考えられるため，このような除外処理は，プローブ情報システム全体のデータ信頼性の向上にもつながると考えられる．これらの軸については，評価軸の整理と客観的な指標の両方について，今後更なる議論が必要となる．

5. おわりに

インターネット等の共通情報通信基盤上に構築されるプローブ情報システムは国内外からの高いイノベーションが期待される一方，車両追跡や行動履歴の推測などの脅威が存在する．各種プローブ情報システムの脅威分析や個人情報保護に関するガイドラインは国内外で検討されているが，プローブデータ自体の信頼性やシステムセキュリティの検証に関する議論は少ない．そこで本論文では，汎用性，社会性の高いプローブ情報システムが扱う情報の信頼性やシステムセキュリティを評価・検証手法を考察した．まず，プローブ情報システムの社会性と ITS 分野における情報セキュリティへの関心の高さについて述べ，プローブ情報システムを実際に実用化するにあたっての事業者の匿名プローブデータへの関心と利用者のデータ発信の不安に関する傾向を掴んだ．利用者は，データの匿名性を説明されるだけでなく，それを検証する手段や自己情報制御機能を必要としている．そのため，プライバシ保護の中で特に匿名性についての一般的な情報セキュリティ評価基準や評価手法を参考にして議論の整理を試みた．その結果，必要な匿名性の強度とその達成度合いという 2 次元の評価軸に整理することが妥当であると云えるが，匿名性強度については個々の要件と汎用性を満たす必要があることから更に議論と調整が必要であることがわかった．また，システム全体の匿名性だけでなく，個々のデータの生起確率が突出している場合は適切な除外処理を行う必要がある．この除外処理は，匿名性確保だけでなくプローブ情報システム全体のデータ信頼性の向上にもつながる．

謝辞

本研究は，経済産業省基準認証研究開発事業「プローブ情報システムにおける個人情報保護に関する標準化」として行われた調査研究を取りまとめたものである．本研究を進めるにあたり，多大なご指導・ご助言をいただいた委員の皆様およびワーキングの皆様に感謝致します．また，日頃の議論や研究活動にご協力いただいた WIDE プロジェクトインターネット自動車 WG の皆様に感謝致します．

参考文献

1) internavi Premium Club インターナビ VICS フローティングカーシステム, http://www.premium-club.jp/technology/tech1.html (2009.8).
2) プローブコミュニケーション交通情報, http://toyota.jp/g-book/mx/technology/probe.html (2009.8)
3) CVIS Project, http://www.cvisproject.org/en/links/xfcd.htm (2009.8)
4) "Vehicle probe data for wide area communications", ISO 22837:2009, http://www.iso.org/iso/catalogue_detail.htm?csnumber=45418 (2009.8)
5) M.SATO, M.IZUMI, H.SUNAHARA, K.UEHARA, J.MURAI: "Threat analysis and protection methods of personal information in vehicle probing system", The Third International Conference on Wireless and Mobile Communications (ICWMC 2007),(2007)
6) M.IZUMI, M.SATO, H.SUNAHARA: "Requirements for protection methods of personal information in vehicle probing system", The 2007 International Symposium on Applications and the Internet(SAINT2007), (2007).
7) ISO/TC204/WG16 HP, http://www.isotc204wg16.org/ , (2009.8)
8) "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data", http://www.oecd.org/document/18/0,3343,en_2649_34255_1815186_1_1_1_1,00.html, (2009/8).
9) "Privacy - the basic principles for probe personal data protection", ISO/DIS 24100, http://www.iso.org/iso/catalogue_detail.htm?csnumber=42017 , (2009.8)
10) Baik Hoh, Marco Gruteser, Hui Xiong, and Ansaf Alrabady: “Preserving Privacy in GPS Traces via Uncertainty-Aware Path Cloaking”, In proceedings of the 14th ACM conference on Computer and communications security, (2007)
11) Bugra Gedik, and Ling Liu: “Location Privacy in Mobile Systems: A Personalized Anonymization Model”, In proceedings of the 25th IEEE International Conference on Distributed Computing Systems (ICDCS’05), (2005).
12) 慶應義塾大学SFC研究所: 平成18年度経済産業省委託事業成果 基準認証研究開発事業(プローブ情報システムにおける個人情報保護に関する標準化) 成果報告書, pp.126-141, (2007).
13) Information technology -- Security techniques -- Evaluation criteria for IT security -- Part 3: Security assurance components: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=46413, (2009.8).
14) Information technology -- Security techniques -- Evaluation criteria for IT security -- Part 1: Introduction and general model: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=40612, (2009.8).
15) Information technology -- Security techniques -- Evaluation criteria for IT security -- Part 2: Security functional components: http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=46414, (2009.8)
16) David L. Chaum, “Untraceable Electronic Mail Return Address, and Digital Pseudonyms”, Communications of the ACM, Volume 24, Number 2, Feb. 1981, (1981).
17) Alastair R. Beresford and Frank Stajano, ”Location Privacy in pervasive Computing”, the IEEE Pervasive Computing, 2003, (2003).
18) 佐藤雅明, 繁富利恵他, “プローブ情報システムのためのプライバシを考慮した匿名認証方式の提案と評価”, 情報処理学会論文誌「新しい時代を切り拓くモバイル通信と高度交通システム」特集, Jan. 2009,(2009).
19) Sk. Md. Mizanur Rahman, Nidal Nasser, et al, “Anonymous authentication and secure communication protocol for wireless mobile adhoc networks”, Security Communication Networks, 2008; 1:179-189, (2008).
20) Andreas Pfitzmann, Marit Hansen, “Anonymity, Unlinkability, Undetectablity, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology—“, Version v0.31, http://dud.inf.ut-dresden.de/Anon_Terminology.shtml , Feb. 15, 2008, (2008).
21) Sebastian Clauss, Stefan Schiffner, “Structuring Anonymity Metrics”, In Proceedings of ACM CCS2006 Workshop on Digital Identity Management (DIM’06), 2006, (2006).