

ビジネスにおけるプライバシー保護技術



日本アイ・ビー・エム株式会社
東京基礎研究所
ID・プライバシー技術担当

沼尾 雅之

Masayuki Numao

Group Leader,
ID & Privacy Technology
IBM Research,
Tokyo Research Laboratory
IBM Japan, Ltd.

プライバシー保護技術は、個人情報をおいかに有効利用したいかという利用モデルが前提にあります。個人情報を一切出さなければプライバシーは保護されますが、それでは、パーソナライゼーションなどの顧客サービスも受けられないわけです。個人情報を利用するITモデルには、「P2P (Peer to Peer)モデル」「コミュニティ・モデル」「クライアント/サーバー・モデル」「企業システム」などがあり、それぞれのモデルにおけるプライバシー保護の技術が開発されてきました。例えば、クライアント/サーバー・システムにおいては、ブラウザを介して、新製品開発などに際して消費者動向を探るためのアンケート調査があり、その分析にはデータ・マイニング技術がよく用いられます。データ・マイニングでは、傾向分析など一般的に成り立つルールを見つけ出すことを目的にしており、この一般化されたルール自体を営業目的に利用することにプライバシー上の問題はありせん。しかし、アンケート収集の過程で個人情報が一度マイニング・エンジンに送られることから、その過程での個人情報の漏えいを危ぐする人もいます。そこでIBMでは、個人の属性にかかわる数値を直接集めず、いったんそれらの数値に乱数を加えてランダム化し、このランダムな数値を多数集めたものから元の属性値の分布を復元する方法を提案しています。また今後は、企業システムにおけるCRM (Customer Relationship Management) ソリューションにも、社員に一律に顧客データのアクセス権限を与えるのではなく、Need-to-Knowを考慮した細かいアクセス制御が必要になります。そこで開発されたのが、Tivoli® Privacy ManagerとTRLプライバシー・アクセス・モニター。短期間かつ低予算で、既存のCRMシステムをプライバシー・ポリシー準拠に変換できます。

Management Forefront 5

SPECIAL ISSUE: Information Security and Privacy

Technologies to Protect Privacy in Business Operations

Techniques for privacy protection depend on models of how effectively a company wishes to use its personal information. Obviously, if no personal information is disclosed, privacy is protected. But without access to personal information a company cannot personalize its customer services.

Privacy protection technology has been developed for a variety of IT models using personal information: the P2P (Peer to Peer) model, community model, client/server model, and business system model. In the case of the client/server model, for example, surveys are conducted with browsers to find consumer trends prior to new product development, and data mining technology is used to analyze the collected data.

The aim of data mining is to conduct trend analyses by discovering general rules. When used only for sales and promotion purposes by the company these generalized rules do not cause privacy problems, but in the course of collecting data, personal information is necessarily transmitted to a mining engine, and this is where concern about information leaks arises.

IBM proposes to eliminate this concern in a way that figures on individuals are randomized by adding random value and then restored to the distribution of the original figures.

Also according to the CRM (Customer Relationship Management) solution, instead of giving all employees access rights to customer information, it is thought that a fine-grained access control -- on a need-to-know basis -- is necessary. The Tivoli® Privacy Manager and the TRL Privacy Access Monitor (Application Privacy Monitoring for JDBC) have been developed for this purpose. With these software, it is possible to make existing CRM systems compliant with a privacy mechanism in a short period of time at low cost.

プライバシー保護技術がなぜ必要か

プライバシーとは何でしょう？ これまで長い間、プライバシーとは「1人にしてもらう権利」のことでした。それがやがて、IT(Information Technology: 情報技術)が発展し、インターネットなどを通じて膨大な個人情報企業が企業に集積されるにつれて、プライバシーの意味が法律的に、またITの見地からも変わってきました。現代のプライバシーとは、「個人、グループ、または組織が、自己に関する情報を、いつ、どのように、どの程度伝えるかを自ら決定できる権利」です。

プライバシーを守る究極の方法は、自分の情報を一切外部に出さないことです。しかし、情報というのは金銭と同じように、必要な相手に伝え、世の中に流通利用されて、初めて価値を持つものです。そこで、プライバシー保護のための規定や、個人情報の管理と利用に関する技術が必要になってきます。

プライバシー保護のための規定には、1980年に発表されたOECD(Organization for Economic Cooperation and Development: 経済協力開発機構)によるガイドラインの「8原則」があり(表1)、その後の法制化や企業がプライバシー・ポリシーを作成する際の準拠とされてきました。

表1. OECDガイドライン8原則

収集制限の原則(Collection Limitation Principle) 法律に従って公正な方法で、かつ(当てはまる場合には)本人の同意を得て情報を収集する
データ内容の原則(Data Quality Principle) データは使用する目的と内容に照らし合わせて、正確・完全・最新であるべき
目的明確化の原則(Purpose Specification Principle) 収集時に目的を明示し、目的外には使用しない
利用制限の原則(Use Limitation Principle) 本人の同意、あるいは法律の要請がある場合以外は、目的外使用しない
安全保護の原則(Security Safeguards Principle) データの滅失、不正アクセス、不正使用、破壊、修正、漏洩防止策をとる
公開の原則(Openness Principle) 個人データ取り扱いのポリシーを公開すること
個人参加の原則(Individual Participation Principle) 本人は、情報の存在の有無を知る権利があり、内容の削除・修正を要求できる
責任の原則(Accountability Principle) データ管理者は、以上の原則に責任を持つ

『プライバシー保護と個人データの国際流通についてのガイドラインに関する理事会勧告』より

さらにヨーロッパでは1995年にEU個人データ保護指令が出されています。これは「諸外国の個人データ保護の水準が適切でない場合、加盟国からのデータ移転を規制する」という厳しい内容のもので、こうした世界情勢に対応することもあって、国内でも遅ればせながら「個人情報保護関連5法」が2005年4月に施行されます。これにより、個人情報のコントロール権は本人に属することが保障され、企業が個人情報を取り扱う場合は、利用の目的を公表し、その範囲内でのみ使用することなどが定められました。

例えば、従来ときとして行われていた「名簿業者からのデータの使用」は、個人情報保護法の下ではできなくなります。また、コール・センターなどで個人識別情報の確認・問い合わせをする場合も、利用目的以外の確認・問い合わせはできなくなります。利用目的以外の確認などをする場合は、その旨本人に知らせ承諾を得なければなりません。

このような新しい状況に対して企業は、営業・企画担当者やコール・センターのオペレーターへのプライバシー教育を徹底させる必要がありますが、それだけでは不十分です。人間系のセキュリティ&プライバシー対策にはどうしても漏れが生じることが多いからです。そこで、システムの中にポリシーに準拠したプライバシー対策を組み込んでおく、新しい考え方と技術が必要です。それがプライバシー保護技術です。

P2Pの情報交換を保護

個人情報を利用するITモデルには、「P2P(Peer to Peer)モデル」「コミュニティー・モデル」「クライアント/サーバー・モデル」「企業システム」などがあります。

大型コンピュータが中心になった時代では、複雑な個人情報処理がほとんどなかったこともあり、プライバシー保護技術としては比較的簡単な認証やアクセス制御だけで済んでいました。しかし、PC(Personal Computer)やインターネットが普及するにつれて、PCなどの端末同士を対等な立場でネットワークに接続し、情報の交換も個人同士で直接行われるような形

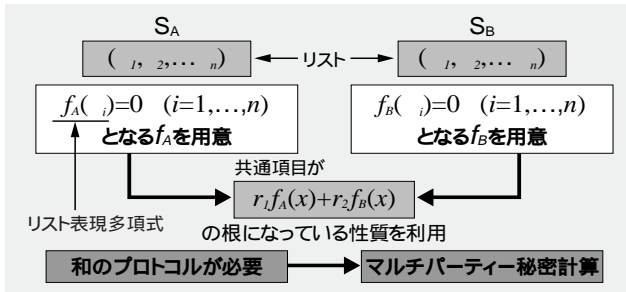


図1. リスト表現多項式の利用

態があらわれました。これが「P2Pモデル」です。

個人同士の情報交換の場合、お互いによく知った間柄であったり、間に調停者(例えば、プロバイダー)がいたりすれば事は簡単なのですが、インターネットでまったく初めて出会った人同士はどのようにすればいいのでしょうか。例えばオークションでは、お金を送ったのに商品が届かないなどの不正がときどき起こっています。このようなことを防ぐには、あらかじめ信頼関係を確立する方法や、暗号技術を利用したプロトコル(通信手順)によって、たとえ不正が起こっても当初の要件が守られるような方法が考えられています。

この問題の解決は、長い間、研究者のチャレンジ対象でした。そして、幾つかの成果を得ています。その一つが「プライバシーを保持したリスト・マッチング」で、リスト表現多項式という多項式表現によって、「マルチパーティー秘密計算」や「忘却多項式評価」という暗号技術を使って、“参加者は共通部分だけを知る”という機密性と、“参加者は、全員が同時に知るか、知らないかのどちらかである”という公平性という要件を満たしています(図1)。

IBMのTRL(Tokyo Research Laboratory:東京基礎研究所)では、例えば、2人以上の参加者が持つリストの共通部分だけを、個々の参加者が持つリストの内容を明らかにすることなく計算できるプライベート・リスト・マッチング・システムを開発しました。これを応用することで、共通の趣味を持つ会員同士を紹介するマッチ・メイキング(お見合い)サービスが可能となります。またビジネスにおいても、複数の金融機関の持つブラックリストの照合や、複数企業の顧客データベースの「名寄せ」に応用できます。

仮想IDの発行

「コミュニティー・モデル」は、P2Pモデルの自然な発展形ともいべきものです。サイバー・スペースではさまざまな趣味や生きがいづくりを目的としたコミュニティー・サイトが続々と生まれていますが、そのコミュニティーを部外者から守りたい、あるいは、コミュニティーの中でも自分のプライバシーは他人には知られたくないといった要望が出てきます。その要望にこたえるための技術には、動的グループ鍵の生成、受信者属性によるメッセージ配信、仮想IDによる属性与信プロトコルなどがあります。

個人の持つ属性は、名前・住所・性別・年齢・職業・年収・趣味・好み・宗教・支持政党...など多種多様です。これらの属性は、第三者(行政機関や団体など)の認定による「認定属性」と個人の裁量で決められる「任意属性」に分けることができます。認定属性については、「住民票」「運転免許証」「就業証明書」などにより認定されますが、趣味・好みなどについての認定証は存在しません。

そこで、本人性の認証と属性の認証を分ける方法が、OASIS(Organization for the Advancement of Structured Information Standard)の標準であるSAML(Security Assertion Markup Language)などでも取り入れられています。

これまでの属性認証の問題点は、例えば、年齢だけを証明したい場合でも住所情報などが併せて記載されていることが多く、必要以上の情報を渡してしまうことです。

「任意属性」は、よりプライベートな属性ではありませんが、同じ趣味を持つ人たちでコミュニティーをつくりたい場合などを考えると、そこに含まれる属性をIT的に利用する意義は決して小さくありません。そのための技術も研究されています。

例えば、金銭以外の信頼情報を与信の対象とし、しかも、取引しようという二者が必ずしも同じ信頼機関に属していない場合、インターネットなどのネットワーク上で属性与信サービスを行うため、IBMでは仮想IDを使った属性認証システムを提案しています(図2)。

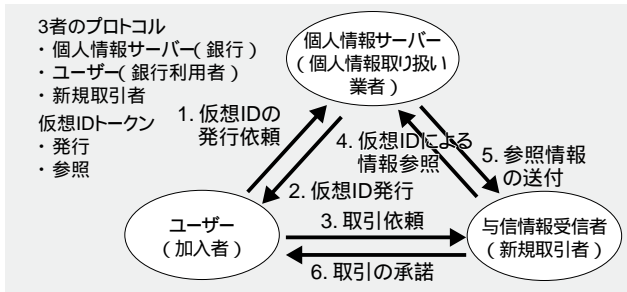


図2. 仮想IDによる属性与信プロトコル

表2. 仮想IDの発行例

GID/VID	名前	口座番号	住所	電話番号	預金残高	...
G000011	鈴木一郎	123456	東京都	333-4444	1,000,000	
V010101	鈴木一郎	-	-	-	1,000,000	
V010011	-	-	東京都	-	1,000,000	

そのシステムの詳細はここでは割愛しますが、仮想IDの発行例を表2に示しておきます。この表のG000011の行がオリジナルの属性レコードです。V010101は残高証明書で、例えば住宅ローン新規借り入れのときの証明書として用いることができます。またV010011は匿名の住所および残高情報だけの統計用データとして用いることができます。このように、顧客レコードから、必要な属性情報だけを証明できるものを、仮想IDによって簡単に実現することができます。

プライバシーを保持したデータ・マイニング

「クライアント/サーバー・モデル」における顧客データベースをビジネスに利用するものに、新製品開発や販売戦略に当たって消費者動向を探るためのアンケート調査があり、その分析にはデータ・マイニング技術がよく用いられます。データ・マイニングでは、購買情報などの個人レベルの情報を大量に集めることによって、関連商品購買分析など、一般的に成り立つルールを見つけ出すことを目的としています。この一般化されたルール自体を公開することにプライバシー上の問題はありませんが、個人情報が一度マイニング・エンジンに送られることから、その過程での個人情報の漏えいを危ぐし、情報提供を拒否する消費者がいるのも事実です。このためIBM Almaden研究所(米

国)では、プライバシーを保持したデータ・マイニング技術を研究開発しました。

データ・マイニングは、給与や年齢などの統計的な分布を知ることが目的ですが、このとき個人の給与や年齢を直接集めずに、いったんそれらの数字に乱数を加えてランダム化します。これを多数集めたものから元の給与や年齢の分布を復元することによって個人のプライバシーを保護します。例えばWebブラウザ上のアンケートの場合、ブラウザのJavaScriptなどによって入力値がランダム化されてサーバーに送られることとなります。また、ランダム化できる対象は数値属性だけでなく、職業などの分散値にも対応可能であり、これにより、プライバシーを保持しながらも精度を損なわずに、関連分析や分類などのデータ・マイニングが可能になっています。

CRM用プライバシー管理システム

「企業システム」におけるプライバシー保護は、経営者にとって最大の関心事でしょう。大量生産・大量宣伝・大量販売というマス・マーケティングに代わって、「顧客一人ひとりにカスタマイズされた製品・サービスを提供し、生涯を通じたロイヤリティーを得る」One to Oneマーケティングが注目され、各企業はこぞって顧客データベースの充実を競ってきたものです。しかし、個人情報保護法の施行以降は、企業がむやみに顧客情

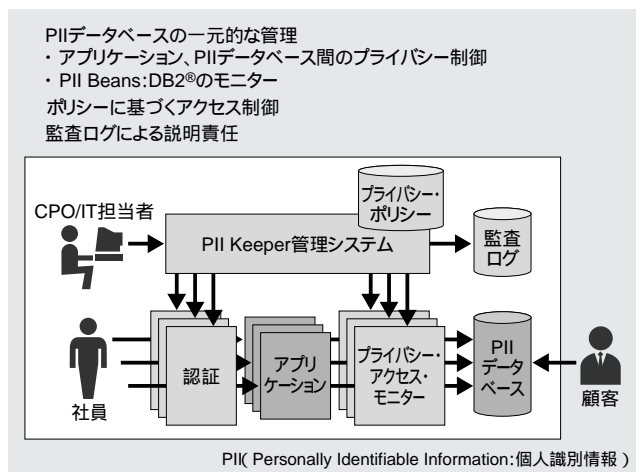


図3. 顧客データ・アクセス管理

報を集めたり利用したりすることができなくなります。顧客データベースを活用するCRM(Customer Relationship Management)ソリューションにも、CRM用プライバシー管理システムが必要になります。

この問題に対してTRLでは、既存のCRMを簡単にプライバシー・ポリシー準拠型システムにするための、プライバシー・アクセス・モニターを開発しました(図3)。

従来のCRMシステムでは、一度システム利用の認証を得たユーザーは、契約社員であろうがパートナー企業の社員であろうと、だれでもがアプリケーションを使って顧客データベースの中に入り込み、このことがさまざまな情報漏えい事件の一因になっていました。そこで、個人情報保護のためには、ユーザーの認証だけでなく、「だれが」「だれの」「どの」「情報に」「どのような目的」でアクセスしようとしているか、という情報をポリシーに照らし合わせて判定する機構が必要であり、また、判定には、所有者の「同意」も必要になります。同意というのは、例えば顧客が自分の情報を登録するときに、表明されたポリシーに同意することであり、この時点で、顧客と企業の間「顧客は個人情報を企業に預ける代わりに、企業はそれをポリシーに準拠して利用する」という契約が成立したとみなされます。また、企業が正しくポリシーに準拠して個人情報を扱っているかの説明責任も企業側にあり、これを満たすために、すべてのアクセス履歴を監査用ログとして保存する義務も生じます。

これらを自動的に行うのが、Tivoli® Privacy ManagerとTRLプライバシー・アクセス・モニターです。Tivoli Privacy Managerはポリシーの生成や管理、ポリシーに基づくアクセスの評価を行うもので、プライバシー・アクセス・モニターは、CRMアプリケーションが

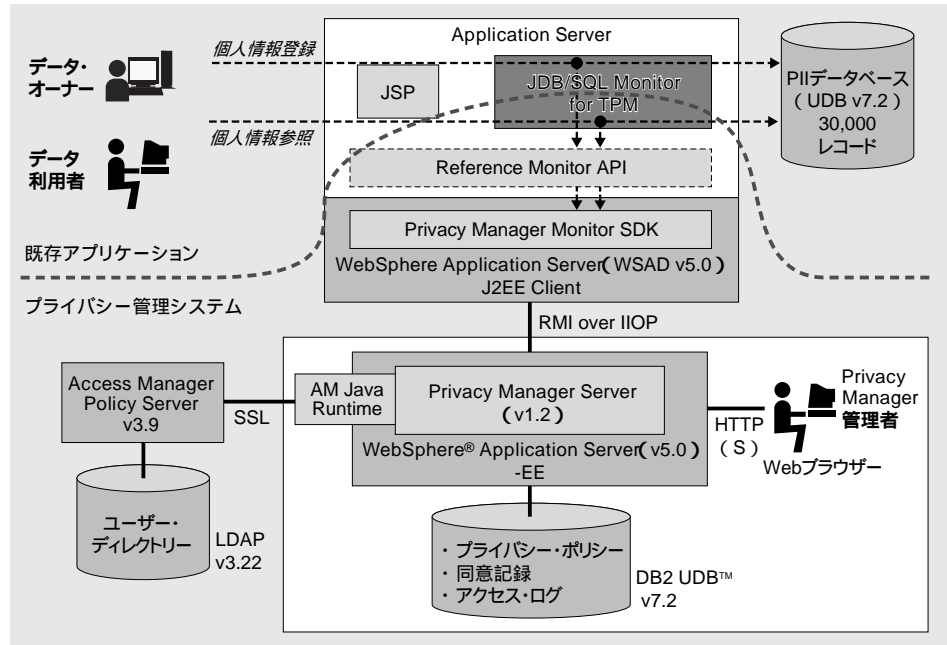


図4. CRM用プライバシー管理システム構成

個人情報データベースをアクセスするときに、SQL (Structured Query Language : 標準照会言語) 文を解析することによって、上記の「だれが」「だれの」「どの」「情報に」「どのような目的」でアクセスしようとしているか、という情報をTivoli Privacy Managerに送って、ポリシーに基づく評価をしてもらい、その結果によって、アプリケーションにどのデータを返すべきかを選定するものです(図4)。

このシステムの特長は、Java™ベースのアプリケーションが通常データベースを呼ぶときに用いるJDBC (Java Database Connectivity) クラスだけをプライバシー・アクセス・モニターに替えれば、既存のアプリケーションやデータベースは一切変えることなく、短期間かつ、低予算で、既存のCRMシステムをプライバシー・ポリシー準拠に変換できることです。

当システム構築の前提には、企業のプライバシー・ポリシーがしっかり確立されていることがあります。その意味で経営者の方々には、ポリシーの重要性を再確認していただき、総合的なセキュリティ&プライバシー対策に取り組んでいただきたいと思います。