センサデータの利活用に適したプライバシ保護付き暗号化システム

井谷　茂寛†　阿比留　健一‡　伊豆　哲也†　伊藤　孝一†
牛田　芽生恵†　小倉　孝夫‡　武仲　正彦†　津田　宏†

株式会社　富士通研究所 † ソフトウェアシステム研究所　セキュアコンピューティング研究部

‡ IT システム研究部　サーバネットワーキング研究部
211-8588 神奈川県川崎市中原区上小田中４－１－１

s.idani@jp.fujitsu.com

あらまし モノ、ヒト、サービスを互いにつなぐ技術として、センサデータの利活用は今後さらに加速すると予想される。一方、スマートメータや位置情報などのセンサデータのプライバシ保護の規制は、EUを中心に厳しくなりつつある。すなわち、センサデータの利活用においては、プライバシ保護と利活用を両立することが今後重要な課題となる。本稿では、プライバシ保護と利活用を両立するためのシステムを考案し、試作を行った。試作システムを用いることで、センサデータを暗号化したまま、高速に部分マスクを行ったり、別の暗号鍵への変換を実現することで、他組織にセンサデータの一部分析を委託する際に必要となる、データの適切なコントロールが可能となる。

Privacy-Preserving Encrypted Data Masking System Suitable for Sensor Data

Shigehiro IDANI†　Kenichi ABIRU‡　Tetsuya IZU†　Kouichi ITOH†
Mebae USHIDA†　Takao OGURA‡　Hiroshi TSUDA†　Masahiko TAKENAKA†

Fujitsu Laboratories Ltd.

†Software Systems Laboratories, Secure Computing Laboratory
‡IT Systems Laboratories, Server Networking Laboratory
4-1-1, Kamikodanaka, Nakahara-Ku, Kawasaki, Kanagawa, 211-8588, JAPAN
s.idani@jp.fujitsu.com

Abstract Use of remotely-sensed data is growing as a technology to connect among humans, devices and services, while rules for privacy-protection are becoming stronger. Therefore, privacy-preserving utilization technique for the sensed data will be a key technology within few years. In this paper, we propose a privacy-preserving encryption technique suitable for the sensed data and show the implementation result of the proposed technique. Our proposed technique can exchange the encryption key to the third party's key and perform dynamic data masking simultaneously without decrypting the sensed data. By using the proposed technique, sensed data can be utilized with given privacy policies even when the data is also utilized by third parties without referring the data not permitted by the policies.

1　はじめに

モノ、ヒト、サービスを互いにつなぐ技術として、センサデータの利活用は今後さらに加速すると予想される。スマートメータの市場規模は 2020 年には 1 兆 5000 億円となると予測されている[7]。また、位置情報の利活用により 6000 億ドルの新規事業の創出効果があると試算されている[11]。

一方、スマートメータや位置情報などのセンサデータのプライバシ保護の規制は、EU を中心に厳しくなりつつある。2012 年 2 月に EU データ保護指令が改版され、2013～2014 年に新たに EU データ保護規則として採択される見込みである。従来の EU データ保護指令では、十分な個人データ保護対策を行っていない第三国へのデータ移転禁止していたたが、今回の改正では、さらに、(a)ICT 技術とグローバル化の進展への対応、(b)多国籍企業への配慮、の2点が追加されている。(a)はクラウドコンピューティングによるまたいだデータ流通や、GPS 等による個人データの収集の高度化に対応するものである。(b)は多国籍企業にとって負担であった、国ごとによるプライバシ保護の基準の不統一を改正するための対応である。従来の EU データ保護指令(Directive)を EU データ保護規則(Regulation)に格上げし、この基準が EU 加盟各国の国内法に優先されるため、統一基準によるプライバシ保護が可能となる。基本的には、オプトイン方式(サービス提供前に、サービス利用者にデータ提供条件を示し同意を得る方式)による自己情報コントロール権を実現することを目的としている。詳細については文献 4] [10]も参照されたい。

スマートメータに関しては、2012 年 6 月に EU データ保護監視官がスマートメータの脆弱性を指摘[13]。同時期に英国のエネルギー・気候変動省はスマートメータのプライバシ規制強化を進めており、本人の合意なしで収集データを第三者提供を禁止する、等の対策を義務化する方針を検討している [9]。

米国では、2012 年 2 月にプライバシ権利章典の草案を公開し、プライバシ保護の規制強化の方針を打ち立てている[14]。米国の場合オプトアウト方式(サービス提供後、サービス利用者の要求に応じてデータの第三者提供禁止等の保護を行う方式)であるが、利用者のリクエストに応じて事業者による利用者の行動追跡防止(Do Not Track, DNT)が加えられていることが特徴である。

日本でも、2012 年 8 月にスマートフォン プライバシ イニシアチブ[8]が総務省によって公表された。現状ではスマートフォンで収集した個人情報のサービス事業者による利活用、利用者に分かりにくい形で使われていることを問題視しており、個人情報をサービス事業者が利活用する場合は、利活用の目的を利用者にわかりやすい形で通知し、同意をとることを目標としている。

国や地域によって方針が異なるが、各国ともプライバシ保護強化の方針を打ち立てており、センサデータの利活用におけるプライバシ保護は今後重要な課題となると予想される。

本稿では、センサデータに適した、プライバシ保護と利活用を両立するための暗号化システムを考案し、試作実装を行った結果を報告する。本稿のセンサデータのプライバシ保護は、プライバシ保護ポリシーに従い、データの提供先に応じて参照可能な情報を制御するために、暗号化方式は以下の 4 つの条件を満たすことを目標とした。(i)センサデータを収集した段階で収集用の鍵を用いた暗号化を行う ことで、不特定多数の情報参照を防ぐ(ii)データ収集を行うサービス事業者は収集用の鍵を有するため全てのデータを参照できる(iii)センサデータの第三者利活用のために、サービス事業者もしくはサービス利用者が設定したプライバシ保護ポリシーに従い、第三者もセンサデータの部分情報を参照可能とする。iv)第三者にセンサデータを提供する場合は、参照不可情報のマスク処理と提供先の暗号化鍵への付け替え処理の両方を行った暗号化データとして提供する。

提案システムを用いることで、センサデータのプライバシ保護と利活用の両立が実現でき、センサデータの分析の外部委託やセンサデータのマーケットプレイスによる第三者利活用も可能にな

る。データ分析を外部委託し第三者利活用を行う場合、分析結果に対するアクセスを行う場合にも、アクセス者の匿名性を実現するプライバシ保護が必要となるが、我々はこの問題についても同様に解決を行っている[1]。

2　提案システムおよび試作結果

本章は、1 章で述べたセンサ向けプライバシ保護を実現するための暗号化技術を用いたシステムを提案する。2.1 章で、センサデータ向けの要件を定義し、要件を満たすプライバシ保護データ利活用システムを提案する。2.2 章でこの要件を満たす暗号方式を説明し、2.3 章では提案システムの試作実装結果について報告を行う。2.4 章では試作実装を用いたシステムデモを紹介する。

2.1　提案システム

第三者を含めた、プライバシ保護付きのセンサデータ利活用を実現するために、本稿では以下をプライバシ保護要件として定義する。

(要件 P-1) センサデータを収集した段階で収集用の鍵を用いた暗号化を行い、センサデータを暗号化して送信することで、不特定多数の情報参照を防ぐ
(要件 P-2) データ収集を行うサービス事業者は収集用の鍵を有するため全てのデータを参照できる
(要件 P-3) センサデータの第三者利活用のために、サービス事業者が設定したプライバシ保護ポリシーに従い、第三者もセンサデータの部分情報を参照可能とする。
(要件 P-4)　第三者にセンサデータを提供する場合は、参照不可情報のマスク処理と提供先の暗号化鍵への付け替え処理の両方を行った暗号化データとして提供する。

すなわち、データ提供者とデータ収集者の二者モデルではなく、第三者も必要に応じてデータを利活用する、という3者モデルを仮定する。

本稿で仮定するプライバシ要件は(要件 P-3)に示されるように、第三者はデータ提供者の全てのデータを参照できるとは限らず、部分参照も可能であることを要件とする。

(要件 P-3)により、第三者に解析を委託する際に、分析に不要な個人情報を第三者に預けることなく、必要な解析処理を行うことができる。

提供者(センサ)、収集者、第三者(解析者)を含んだ、(要件 P-1)～(要件 P-4)を満たす構成を図1に示す。

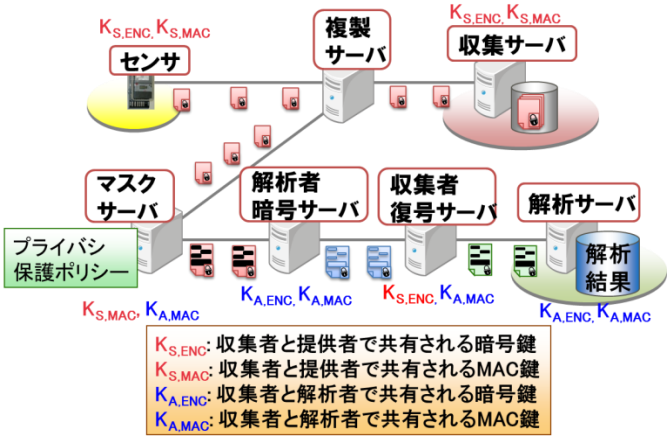

図 1 提案システム全体図

センサと収集者の間では鍵 $K_{s,Enc}$, $K_{s,MAC}$ が共有され、解析者と収集者の間では鍵 $K_{a,Enc}$, $K_{a,MAC}$ が共有される。

センサは、センサデータを自動的に鍵 $K_{s,Enc}$ で暗号化し、鍵 $K_{s,MAC}$ で MAC をつけて収集サーバに送信する。複製サーバは、センサから収集サーバに送られる暗号化されたデータを複製し、収集サーバと解析サーバへ送信する。複製されたパケットは、解析サーバにたどり着く前に、マスクサーバ、解析者暗号サーバと収集者復号サーバを中継する。

マスクサーバで暗号化されたデータの一部を、解析者暗号サーバ毎に用意されている、プライバシ保護ポリシーに応じてマスク化処理を行う。改ざん防止のため、受信データは鍵 $K_{s,MAC}$ で MAC を確認し、送信時には鍵 $K_{a,MAC}$ による MAC をつける。

そして解析者暗号サーバでは、暗号データを鍵 $K_{a,Enc}$ で暗号化し、改ざん防止として鍵 $K_{a,MAC}$ を

使用する。

次に収集者復号サーバでは、受信した暗号データを鍵 $K_{s,Enc}$ で復号を行い、改ざん防止として、鍵 $K_{a,MAC}$ による MAC をつける。

最終的に解析サーバには、受信した暗号データの改ざん防止を鍵 $K_{a,MAC}$ で確認し、鍵 $K_{a,Enc}$ で暗号化されたデータが手に入る。

この構成を用いることで、データを平文に戻さずに暗号化したままマスク化を行い、第三者に、マスク化されたセンサデータの部分情報を提供することが可能となる。

素朴な方法である、データを平文に戻してから、平文に対してセンサデータのマスク化を行い、解析者の鍵で暗号化する方法の場合、平文を扱えるのが収集サーバに限定されるため、大量のセンサデータを取り扱う場合処理負荷が高くなるという問題が発生する。しかし、提案システムでは、データを平文に戻さずに暗号化したままマスク化を行うことで、この問題を回避することができる。

2.2 提案システムに関する暗号の要件

提案システムにおける暗号化処理の要件を定義する。

(要件 1). センサから収集されたデータを収集者のみが利活用を行うという二者モデルではなく、必要に応じてセンサデータの一部解析を第三者(解析者)に委託する三者モデルであること。

(要件 2). センサはデータを収集した後、収集者の鍵で暗号化を行い収集者に送信すること。

(要件 3). 収集者がセンサから受信した暗号化センサデータの解析を解析者に委託する場合は、以下の3つの条件を満たした暗号化処理を実現する。

A) 収集者の鍵による暗号化データを解析者の鍵に付け替える。

B) 収集者が解析者に見せたくない属性については、マスク処理によって隠す。

C) 上記の鍵付け替えとマスク処理は、センサデータを復号することなく実現する。

(要件 4). 上記の暗号化処理はセンサのような計算処理が少ない環境でも実現可能であること。

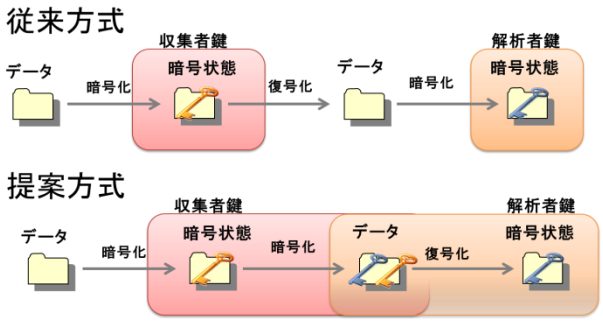

図 2 従来と提案手法図

従来の暗号文の鍵の付け替えは、図 2 の「従来方式」に示すように、いったん暗号文を復号し、平文に戻し、その上で付け替えたい鍵で暗号化することで実現していた。

しかし、この場合は、要件 3-C を満たさない。

また、文献[12]も鍵付け替え可能な暗号について提案を行っているが、公開鍵方式であるため、処理が遅く、要件 4 を満たさない。先の要件を満たすには、図 2 の「提案方式」に示すように、収集者の鍵による復号と、解析者の暗号鍵での暗号化で「可換性」を満たす必要がある。

任意の平文を M と任意の鍵を K_1、K_2 に対して、数式 1 が成立する。

数式 1

$$E_{K2}\{ D_{K1}(M) \} = D_{K1} \{ E_{K2} (M) \}$$

この「可換性」の性質をもつ暗号方式を適用することで、先の暗号要件を満たすことが可能となる。この時に以下が成立する。

数式 2

$$E_{K1}[E_{K2}\{ D_{K1}(M) \}]$$
$$= E_{K1} [D_{K1}\{ E_{K2}(M) \}]=E_{K2}(M)$$

次に「可換性」を持つ暗号アルゴリズムを検討する。

通常の暗号アルゴリズムは、可換性をもたない。しかし、多くのストリーム暗号では、数式 3で実現がされているため、可換性を持つ。

数式 3

$$Enc_{鍵}(M) = M \oplus 疑似乱数(鍵)$$

これを示したのが図 3 である。

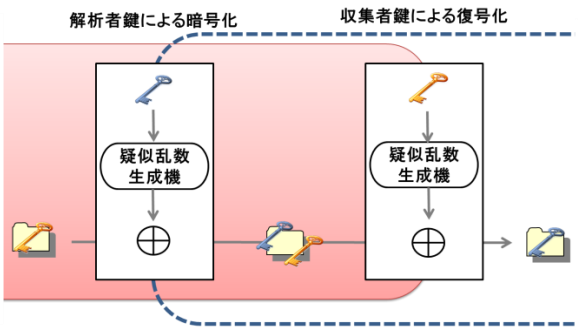

図 3 共通鍵ストリーム暗号

また、共通鍵暗号のカウンタ(CTR) モードは、数式 4によって実現されている。こちらも「可換性」をもつと言える。

数式 4

$$Enc_{鍵}(M) = M \oplus 共通鍵暗号_{鍵}(CTR)$$

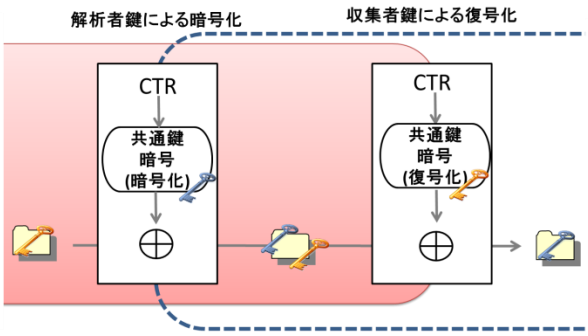

図 4 共通鍵カウンターモード

以上の 2 方式が可換性を持ち、先のシステム要件を満たす暗号アルゴリズムと言える。

この暗号アルゴリズムを提案システムに適用することで、プライバシを保護したデータ利活用システムが実現可能であることを示す。

図 1のセンサが、収集したデータに含まれる x 個の各属性を M_1、M_2、M_3、・・・、M_x とする。収集サーバとの共有鍵($K_{s,Enc}$)とする。この暗号化されたデータを数式 5に示す。

数式 5

$$E_{Ks,Enc}(M_1) \,||\, E_{Ks,Enc}(M_2) \,||\, E_{Ks,Enc}(M_3) \,||\, \cdots \,||\, E_{Ks,Enc}(M_x)$$

暗号化されたデータは、複製サーバで中継、複製され、元々の送信先である収集サーバと解析サーバへデータに送られる。

マスクサーバでは、マスクサーバのみが有するマスク用の鍵($K_{m,Enc}$)で、解析サーバには見せたくないデータをマスク化する。ここでは属性 M_2 をマスク化するものとすると、数式 5は数式 6となる。

数式 6

$$E_{Ks,Enc}(M_1) \,||\, E_{Km,Enc}\{E_{Ks,Enc}(M_2)\} \,||\, E_{Ks,Enc}(M_3) \,||\, \cdots \,||\, E_{Ks,Enc}(M_x)$$

解析者暗号サーバは、解析サーバと鍵($K_{a,Enc}$)を共有し、受け取ったデータを鍵($K_{a,Enc}$)で暗号化する。これにより数式 6は数式 7となる。

数式 7

$$E_{Ka,Enc}\{E_{Ks,Enc}(M_1)\} \,||\, E_{Ka,Enc}[E_{Km,Enc}\{E_{Ks,Enc}(M_2)\}] \,||\, E_{Ka,Enc}\{E_{Ks,Enc}(M_3)\} \,||\, \cdots \,||\, E_{Ka,Enc}\{E_{Ks,Enc}(M_x)\}$$

次に収集者復号サーバは、収集サーバの鍵($K_{s,Enc}$)を持ち、復号処理を行う。

これによって、数式 7は、数式 8となる。

数式 8

$$D_{Ks,Enc}[E_{Ka,Enc}\{E_{Ks,Enc}(M_1)\}] \,||\, D_{Ks,Enc}[E_{Ka,Enc}[E_{Km,Enc}\{E_{Ks,Enc}(M_2)\}]] \,||\, D_{Ks,Enc}[E_{Ka,Enc}\{E_{Ks,Enc}(M_3)\}] \,||\, \cdots \,||\, D_{Ks,Enc}[E_{Ka,Enc}\{E_{Ks,Enc}(M_x)\}]$$

今回使用した暗号アルゴリズムは可換性をもつため、数式 8は数式 9に変換ができる。

数式 9

$$E_{Ka,Enc}[D_{Ks,Enc}\{E_{Ks,Enc}(M_1)\}] \,||\, E_{Ka,Enc}[E_{Km,Enc}[D_{Ks,Enc}\{E_{Ks,Enc}(M_2)\}]] \,||\, E_{Ka,Enc}[D_{Ks,Enc}\{E_{Ks,Enc}(M_3)\}] \,||\, \cdots \,||\, E_{Ka,Enc}[D_{Ks,Enc}\{E_{Ks,Enc}(M_x)\}]$$

数式 9は、数式 2のように処理ができるため、数式 10となる。

数式 10

$$E_{Ka,Enc}(M_1) \,||\, E_{Ka,Enc}\{E_{Km,Enc}(M_2)\} \,||\, E_{Ka,Enc}(M_3) \,||\, \cdots \,||\, E_{Ka,Enc}(M_x)$$

最後に解析サーバが数式 10のデータを受け取り、鍵($K_{a,Enc}$)で復号することで、マスク化された M_2 以外のデータを得ることが可能となる。

2.3　提案システムの試作

本提案システムを試作するに辺り、データを中継・処理する、「マスクサーバ」、「解析者暗号サーバ」、「収集者復号サーバ」には、文献[2]、[3][5][6]で紹介している『情報ゲートウェイ』を適用した。

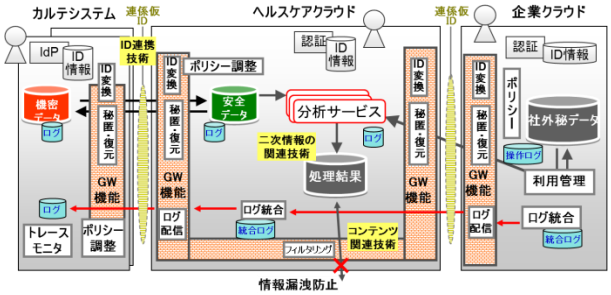

図 5 情報ゲートウェイ全体図

『情報ゲートウェイ』は、図 5 に示すように、ID変換技術、秘匿・復元技術、ログ配信技術、フィルタリング技術、サービス提供技術、トレース技術など、多様なサービスに対応するよう設計された、サーバソフトウェアである。

その後、中継するデータを加工し、特に秘匿・復元機能に特化した個人端末向けの、『パーソナル情報ゲートウェイ』の開発を行った。

パーソナル情報 GW の機能

『パーソナル情報ゲートウェイ』は、Windows PC や Android OS 上でも動作するために、以下の要件を設定した。

- さまざまな OS で動作すること
- さまざまな Web クライアントで動作すること
- 秘匿復元に関して拡張性があること

この要件を満たすために、OS 上で動作する Proxy アプリケーションとして実装を行った。
また、拡張性を確保するため、プラグイン機能を実装した。

試作したパーソナル情報ゲートウェイは、特定のローカルポートで待ち受けを行い、以下の3つの項目に対しパターンマッチで秘匿・復元処理を行う

- URL
- HTTP ヘッダー
- HTTP ボディ

Google カレンダーを例にすると、カレンダーに新規予定を登録する Post 処理で、HTTP のヘッダーとボディに書かれている予定を秘匿化した予定に置き換える。また、Get 処理で取得する予定に対して、自動的に復元処理を行う。これにより、Google のクラウドサービスに対し、スケジュール情報を秘匿化して預ける等の機能を実現できる。

試作で定義したカスタマイズ内容

センサ情報を秘匿化するにあたって、「パーソナル情報ゲートウェイ」のプラグイン機能で以下の 3 種類の実装を行った。

1. 解析者鍵による暗号プラグイン
2. 収集者鍵による復号プラグイン
3. マスク化プラグイン

実装は Java で行い、暗号処理には、JAVA 暗号化拡張機能(JCE)を利用し、暗号方式は、鍵長 128bit の AES のカウンターモードを採用した。

センサデータを扱う関係上、固定長データを想定していたが、運用において可変長も取り扱う可能性がでてきた。このためカウンタを意識しなくてもよいように、データごとに IV を新規に生成する方式を採用した。

2.4　試作したデモシステム

図 1をもとにデモシステムを構築した。

通信データを中継し、再暗号を行う処理は、先に説明した「パーソナル情報ゲートウェイ」で試作した。ここでは、その他の「センサ端末」、「解析サーバ」、「収集サーバ」の説明を行う。

「センサ端末」は、Java で実装を行い、指定した時間で疑似データを生成し送信する。生成するデータは、水道、ガス、電気のメータと仮定し、以下の項目とした。

- 顧客 ID
- メータ ID
- 検針値
- 検針日時

この 4 項目を暗号化し、それぞれの項目の暗号化に使用した IV を先頭に追加して、図 6の形式で送信をする。

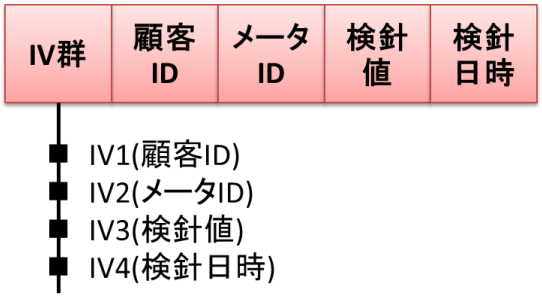

図 6送信データフォーマット

収集サーバと解析サーバは、Java で実装を行った。図 7がデータを受信する画面である。

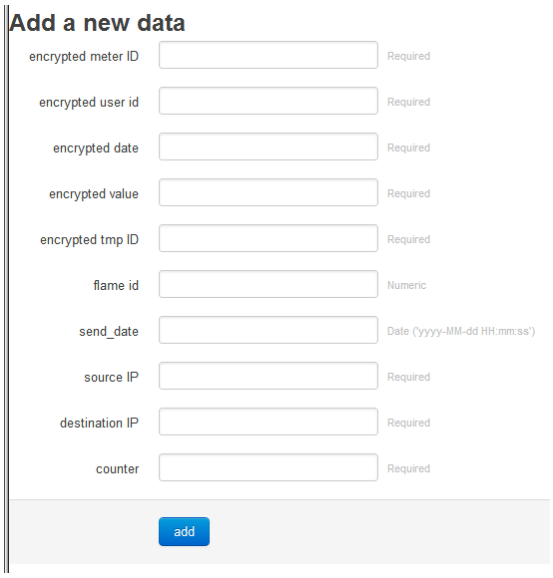

図 7収集・解析サーバ データ受信画面

受信したデータは暗号化したまま MySQL に登録する。また、図 8のように受信データを復号し、閲覧も可能である。ただし、解析サーバは、一部データがマスク化されているので、図 9のようにマスクされた状態で表示される。

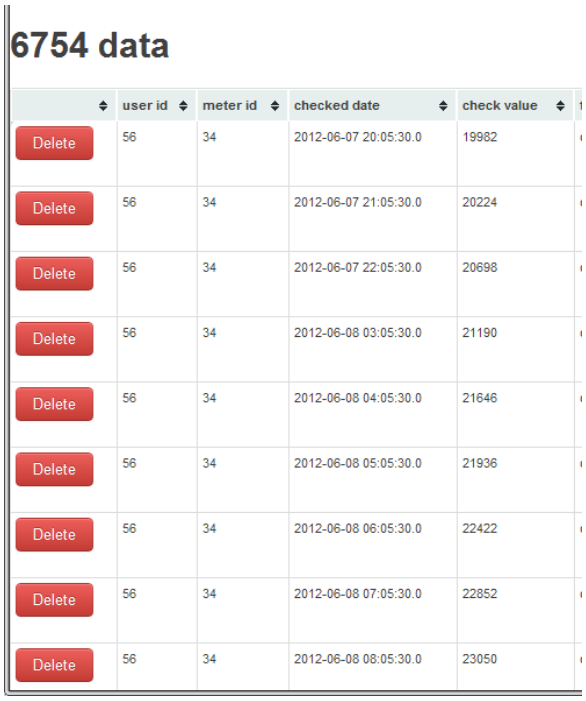

図 8 収集サーバ 受信データ表示画面

図 9解析サーバ 受信データ 表示画面

図 10 デモシステム ログ

デモシステムは、デモ性質上、途中の暗号処理も見せるため、各データを Syslog サーバに送る機能を追加した。この表示データが図 10である。

マスクサーバで 1 列目のデータの「3ce3」から「57e9」に暗号化されマスク化される。その後、

「解析暗号サーバ」、「収集復号サーバ」で暗号処理と復号処理が行われることで、暗号鍵の付け替えが行われる。最終的に受信する解析サーバでは、マスク化された 1 列目以外が復号される。

3 まとめ

センサデータに適した、プライバシ保護と利活用を両立するための暗号化方システムを考案し、試作を行った。提案システムを用いることで、センサデータのプライバシ保護と利活用の両立が実現でき、センサデータのマーケットプレイスや分析の外部委託等による第三者利活用も可能になる。

参考文献

[1] 伊豆, 阿比留, 井谷, 伊藤, 牛田, 小倉, 武仲, 津田, “プライバシ保護を考慮したセンサデータの利活用について”, コンピュータセキュリティシンポジウム, 2012.10

[2]牛田 芽生恵, 伊藤 孝一, 片山 佳則, 小櫻 文彦, 津田 宏, ゲートウェイによるクラウド間のデータ秘匿集計技術, 暗号と情報セキュリティシンポジウム(SCIS2011), 2011.1

[3] 片山,高,小櫻,津田, クラウドにおけるデータ秘匿・追跡技術とその応用, コンピュータセキュリティシンポジウム, 2010.10

[4] 小泉 雄介, "グローバルな個人データ保護制度改定の動向と日本企業への影響", http://www.i-ise.com/jp/study/EUdata_20120725.pdf

[5]津田, クラウド連携における個人・機密情報の利活用, 情報処理学会連続セミナー「ビッグデータとスマートな社会」, 2012.7.17

[6]津田,伊藤, プライバシーを含む Linked Data の秘匿分析技術, 人工知能学会全国大会, 2012.6

[7] 日経 BP 社, “世界のスマートメーター市場規模は 2020 年に 1 兆 1500 億円”, http://corporate.nikkeibp.co.jp/info/newsrelease/newsrelease20110725.shtml

[8] 利用者視点を踏まえた ICT サービスに係る諸問題に関する研究会, “スマートフォン プライバシー イニシアチブ”, http://www.soumu.go.jp/main_content/000171225.pdf

[9] Depertment of Energy & Climatechange “Smart Metering Implementation Programme”, http://www.decc.gov.uk/assets/decc/11/consultation/smart-meters-security-risk-assess/5434-cons-smart-meters-security-risk-assess.pdf

[10] JEITA, “EU データ保護指令改定に関する調査・分析報告書”, http://home.jeita.or.jp/page_file/20120427161714_ljwGedIUnB.pdf

[11] McKinsey Global Institute, “Big data: The next frontier for innovation, competition, and productivity”, http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation

[12] Ryotaro Hayashi, Tatsuyuki Matsushita, Takuya Yoshida, Yoshihiro Fujii, and Koji Okada , “Unforgeability of Re-Encryption Keys against Collusion Attack in Proxy Re-Encryption”, Advances in Information and Computer Security – IWSEC 2011, LNCS 7038, pp.210-219

[13] The register, “Smart meters are 'massive surveillance' tech - privacy supremo”,http://www.theregister.co.uk/2012/06/11/smart_meter_privacy/

[14] The White House “Consumer Data Privacy In a Networked World: A Framework For Protecting Privacy and Promoting Innnovation in the Global Digital Economy”, http://www.whitehouse.gov/sites/default/files/privacy-final.pdf