プライバシー保護を考慮したセンサーデータの利活用について

伊豆 哲也† 阿比留 健一‡ 井谷 茂寛† 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
†セキュアコンピューティング研究部 ‡IT システム研究所
§基盤ネットワーキング研究部 ¶ソフトウェアシステム研究所
〒 211-8588 神奈川県川崎市中原区上小田中 4-1-1
izu@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, センサーデータの利活用が広がっている. しかし欧州では 2012 年 2 月にデータ保護指令が改版されるなど, 個人データに対するプライバシー保護規制が厳格化されており, センサーデータの利活用において, 可用性を低下させることなくプライバシー保護を実現させることが課題となっている. 本稿では, CSS 2012 において井谷等が提案したマスク化可能暗号方式に着目し, 消費電力センサーを用いたエネルギーマネジメントシステム (EMS) に適用した場合の活用法と課題について検討する. 特に, 消費電力解析者に対するプライバシー保護手段として, アクセスゲートウェイを用いた仮名化手法を提案する. また, フローティングカーデータシステム (FCDS) への適用もあわせて検討する.
キーワード センサーデータ, プライバシー保護, マスク化可能暗号

Privay-Preserving Utilization of Sensor Data

Tetsuya Izu† Kenichi Abiru‡ Shigehiro Idani† Koichi Itoh†
Mebae Ushida† Takao Ogura§ Masahiko Takenaka¶ Hiroshi Tsuda¶

FUJITSU LABORATORIES Ltd.,
†Secure Computing Lab. ‡IT System Laboratories
§Server Networking Lab. ¶Software Systems Laboratories
4-1-1 Kamikodanaka Nakahara-ku, Kawasaki, 211-8588, Japan
izu@jp.fujitsu.com

Abstract Utilization of sensor data is paid attention since it connects things, humans and services to each other. Because of the privacy reasons, the availability and privacy-preserving property should be compatible when sensor data is used. In this article, we consider practical applications of the encrypted data masking technique proposed by Idani et al. at CSS 2012 to the energy management system (EMS). Especially, in order to establish the privacy to the analyst, we propose the anonymization technique via the access gate-way. We also discuss the application to the floating car-data system (FCDS).
Key Words Sensor data, privacy-preserving, encrypted data masking

1　はじめに

モノ・ヒト・サービスを互いにつなぎあい, 新たな価値を生み出す源として, センサーデータが注目を集めている. 例えば, スマートメータを用いたスマートグリッドや EMS (エネルギーマネジメントシステム), 温度・湿度・照度などのセンサーデータを活用した農業支援システム, 建設機械・製造機械のメンテナンス情報を遠隔で一括に収集管理するシステム, 位置情報や加速度センサー情報によるフローティングカーデータシステム (通行実績サービス, 渋滞予測サービス) など, 新しい利活用がさまざまな分野で広がりつつある. また, データマーケットプレースを通じてこのようなデータを結びつけあうことで, さらなる価値の発見を目指す動きも見られる. 実際, センサーデータビジネスの成長性は有望視されており, 全世界のセンサーデバイス市場規模は 2010 年の 3 兆 379 億円が 2020 年には 4 兆 5293 億円 (49%増) に, 国内のセンサーネットワークソリューション市場規模は 2010 年の 1112 億円が 2020 年には 2632 億円 (136%増) に成長すると見込まれている [2].

このような社会の変化に対応するため, 欧州のデータ保護指令が 2012 年 2 月に改版された. 従来のデータ保護指令では, 十分なデータ保護対策を行っていない第三国へのデータ移転が禁止されていたが, 国をまたいだクラウドサービスによるデータ流通を許容するなど, データ利用の可用性を向上させる変更がなされている. プライバシー保護の観点では, 忘れ去られる権利 (Right to be Forgotten, データ保持者に要求することでユーザが自己に関するデータを削除できる権利) を含む自己情報のコントロール権の確立を目指しており, データ保護指令として, 個人データ利用における可用性とプライバシー保護の両立を求めている. 新しいデータ保護指令は 2 年間のレビューを経た後, 2014 年以降に EU データ保護規則として採択される見込みであるが, 1995 年に制定された (以前の) データ保護指令が日本の個人情報保護法制定に影響を与えたことを考慮すれば, 新しいデータ保護規制も日本に再び影響を与えることは容易に想像できる.

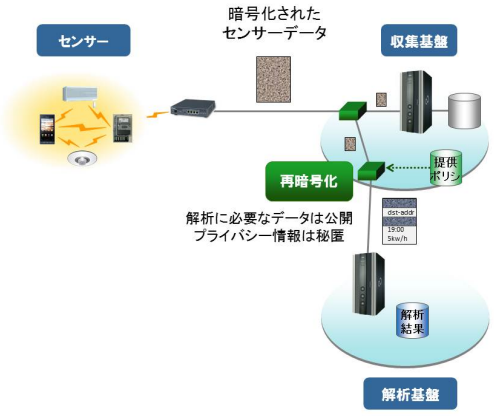

図 1: 再暗号化方式の構成

本稿の目的は, センサーデータのような大量のストリームデータを利活用する場面で, 可用性とプライバシー保護とを両立させる手段について検討・考察することである. プライバシー保護を考慮したデータ活用技術として, プライバシー保護データマイニング (PPDM) [1] や秘匿集計 [5], [11], 完全準同型暗号 [3] などが提案されているが, その多くは処理が複雑であり, ストリームデータへ直接適用することは困難である. そこで本稿は, 井谷等が提案するプライバシー保護を考慮した暗号化方式であるマスク化可能暗号 [4] に着目し, 消費電力センサーを用いたエネルギーマネジメントシステム (EMS) に適用した場合のシナリオと, 課題を抽出・議論する. 特に, 消費電力解析者に対するプライバシー保護手段として, アクセスゲートウェイを用いた仮名化手法を提案する. また, フローティングカーデータシステム (FCDS) への適用もあわせて検討する.

2　マスク化可能暗号

井谷等は CSS 2012 において, ストリームデータのプライバシー保護に適したマスク化可能暗号方式を提案した [4]. この暗号方式では, センサーが観測したデータは暗号化されて収集基盤に送信・収集されるが, センサーデータの第三者利用 (解析基盤におけるキュレーション) を可能とするために, 収集基盤向けの鍵で暗号化

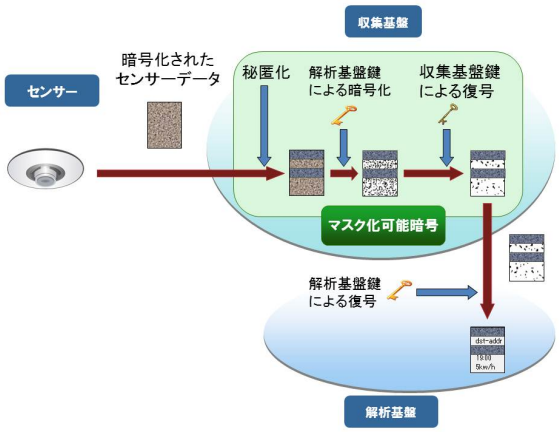

図 2: マスク化可能暗号方式の原理

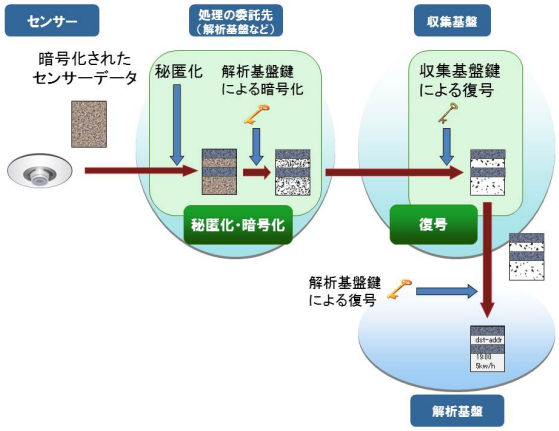

図 3: マスク化可能暗号方式の原理 (分離型)

されたデータは解析基盤向けの鍵で暗号化されたデータに変換 (再暗号化) される．さらに, センサーデータのプライバシー保護を実現するために, 変換時にもとのセンサーデータを経由しない (平文に復号しない) ことや, 解析基盤に秘匿すべき部分情報を秘匿化することを特徴としている (図 2 参照).

暗号化データを解析基盤に提供する場合, 復号してからプライバシー情報を秘匿化 (削除) し, 解析基盤向けに暗号化して提供する素朴な手順が考えられる．しかしこれらの操作は平文情報を対象としているため, 秘密保持の理由から収集基盤しか処理することができず, 収集基盤の負担の増大を招いてしまう．しかしマスク化可能暗号方式では, 秘匿化や解析基盤向けの鍵による暗号化は別の処理基盤 (例えば解析基盤) へ委託することが可能であり, 収集基盤は処理の増大を抑制したままで, 暗号化データの外部提供を実現できる (図 3 参照).

前述したような処理を実現するために, マスク化可能暗号方式は, 任意の平文 M と任意の鍵 K_1, K_2 に対して

$$E_{K_2}(D_{K_1}(M)) = D_{K_1}(E_{K_2}(M))$$

が成立するという “可換性” を必要とする．ここで関数 $E_K(\cdot)$, $D_K(\cdot)$ は鍵 K による暗号化関数と復号関数である．このとき, 任意の平文 M と任意の鍵 K_1, K_2 に対して

$$\begin{aligned} & D_{K_1}(E_{K_2}(E_{K_1}(M))) \\ = \; & E_{K_2}(D_{K_1}(E_{K_1}(M))) \\ = \; & E_{K_2}(M) \end{aligned}$$

となるため, 上記のように平文を経由しない再暗号化が可能となる (秘匿化についても同様). このような可換性を持つ共通鍵暗号の実現方法については, 井谷等の論文 [4] を参照されたい.

3 EMS への適用検討

2011 年 3 月に発生した東日本大震災を発端として, 全国的に電力供給が不安定となっている. 経済産業省は消費電力の可視化と省エネを推進するためにエネルギー管理システム導入促進事業 [6] を開始させており, さまざまな企業がエネルギーマネジメントシステム (EMS) を提供するに至っている．例えば家庭用 EMS (HEMS) では NTT の FLET’S ミルエネ [7], KDDI のエコビトなどが (トライアル) サービス提供を開始している.

本節では, HEMS におけるデータの流れを整理した上でプライバシー問題を抽出し, 再暗号化方式を用いた保護方法を検討する.

3.1 HEMS におけるデータの流れ

多くの HEMS では, クーラーなどの消費電力が大きな家電製品を接続するコンセントや配電盤に消費電力センサーを設置し, 観測した消費

電力情報を収集基盤に送信・収集している．ユーザは，可視化された消費電力情報や解析情報 (過去１年間の消費電力推移に基づく解析結果など) を収集基盤から受け取ることで，節電を推進することが期待されている．

このような HEMS サービスは国内では始まったばかりであるが，海外では先行事例が散見できる．Google は 2009 年 2 月に米国にて PowerMeter という消費電力の可視化・分析サービスの提供を開始した．しかし，ユーザ (および電力会社) は PowerMeter 専用のスマートメータを設置する必要があり，ユーザ数が伸びなかったことから，2 年後の 2011 年 9 月にサービス提供が終了した．Microsoft も 2009 年 6 月に米国で Hohm という消費電力の可視化・解析サービスの提供を開始したが，やはりユーザ数が伸びなかったため，2012 年 5 月にサービス提供を終了している．2007 年に設立された消費電力の解析専門企業 (キュレーター) である Opower は，(PowerMeter や Hohm とは異なり) 電力会社に解析技術を提供することで，ユーザ数の確保に成功している．また，近所の同規模の家庭との比較結果を提供するなど競争の原理を取り入れることで，ユーザが解析情報を有効活用する率が高く，支持を得られている．

3.2　プライバシー問題と対策

現在の HEMS では，観測された消費電力は収集基盤によって解析 (集計) されている．しかし，Opower のような高度なデータ解析を実現するには，第三者による解析は欠かせない．また，米国のように電力供給が自由化されている場合には，解析業者は消費電力推移に基づいた最適な電力供給会社や契約プランを推奨しようとすれば，電力会社とは独立した解析企業の利用が必須である．さらには，スマートシティの実現を目指す地域では，コミュニティ内の家庭・企業の消費電力を把握した上で，地域全体としてのエネルギーの最適化を図るマネジメントシステム (CEMS) を構築するため，各家庭の消費電力情報は CEMS の管理基盤に集約される必要が生じると考えられる．

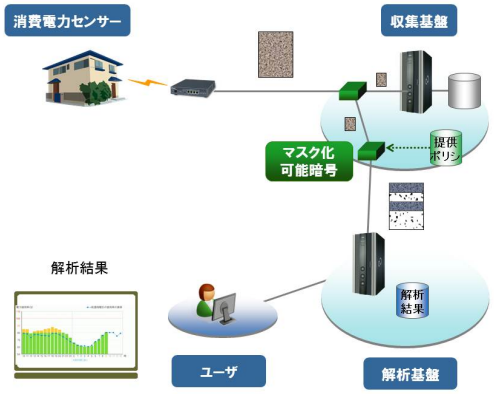

図 4: HEMS におけるプライバシー保護

このように HEMS における情報解析の重要性を考慮すれば，消費電力データは収集基盤に集約される一方で，外部の解析基盤や管理基盤にも送信される必要が生じる．しかし元の消費電力データには収集基盤以外には提供することの出来ないプライバシー情報が含まれていることから，保護手段を導入する必要が生じる．

前節で述べたような HEMS におけるプライバシー保護を実現する上で，井谷等のマスク化可能暗号方式を適用した構成を図 4 に示す．消費電力センサーは消費電力を観測し，収集基盤鍵によって暗号化された消費電力データを収集基盤に送信する．収集基盤は受信した暗号化消費電力データを集約する一方で，暗号化消費電力データを解析基盤に提供するために，不要なプライバシー情報は秘匿した上で，解析基盤鍵によって再暗号化処理を行い，その結果を解析基盤に提供する[1]．そしてユーザは解析基盤から解析結果を受け取ることになる．

このように 図 4 の構成をとることで，HEMS においてデータの第三者 (解析基盤への) 提供が可能となり，暗号化消費電力データの可用性を低下させることなく，プライバシー保護を実現することができる．

[1] プライバシー情報の秘匿及び解析基盤鍵による暗号化処理は収集基盤以外が処理することも可能であるが，ここでは簡単のため，収集基盤が処理する場合を想定する．

3.3　解析者に対するプラバシー問題と対策

前節では, センサーデータが解析者に提供されるまでの過程におけるプライバシー問題を検討した. しかし 図 4 の構成では, 解析者にユーザ ID とセンサーデータの双方が提供されている. そこで本節では, アクセスゲートウェイを導入し, 仮名 ID を用いた匿名化を用いることで, このようなプライバシー問題を解消できることを述べる.

3.3.1　アクセスゲートウェイについて

アクセスゲートウェイ (以下, アクセス GW と呼ぶ) は, サービスとユーザ間をつなぐ技術としてデータを預けたクラウドと連携し，クラウド間のデータアクセス制御を実現する技術である [8, 9, 10]. 企業ユースとして利用する場合には，企業ユーザの属性 (部署，役職など) に応じた利用サービスや預けたデータについてのアクセス権限に応じてアクセス制御を行う. アクセス GW では, 各サービス事業者と契約するたびに，通信プロトコル，認証方式などが異なるとその設定，が大変になるが，このようにアクセスゲートウェイでまとめて制御を行うことで，企業側の対応を軽減できる効果がある. また, ユーザはネットワークサービス（アクセス GW）にログインすれば, アクセス GW で代理認証を行うため, 各サービスへの SSO (Single Sign On) が可能となる.

3.3.2　課題

解析基盤に対するプライバシー保護を確立するために, 収集基盤では, 収集データがどのユーザであるかを特定できないように収集データの一部であるプライバシー情報（ここでは, ユーザを特定する ID や氏名, 住所等）を秘匿化し, 第三者 (解析基盤) に提供することが考えられる. この場合, 解析基盤はユーザを特定せずに解析するのでユーザにとって安心である. また, 解析基盤も, 余計なユーザ管理やセキュリティ保護をしなく済む. その反面, 解析基盤に通知されるデータは, プライバシー情報等が秘匿されているため, 何を識別子として解析したらよいか, わからないという問題が生じる. これを解決するため, 収集基盤と解析基盤の間で仮名 ID を用いる. 具体的には, 収集基盤では, 集めてきたセンサーデータに仮名 ID を付加して解析基盤に提供する. 解析基盤では, 仮名 ID を識別子として, それに基づく収集データの分析を行う. また, 解析基盤では, その分析したデータをユーザへ提供する必要がある. その際, ユーザは解析データを参照する場合も, 解析基盤に特定されることなく, 解析データを取得できる必要があり, これを実現することが課題である.

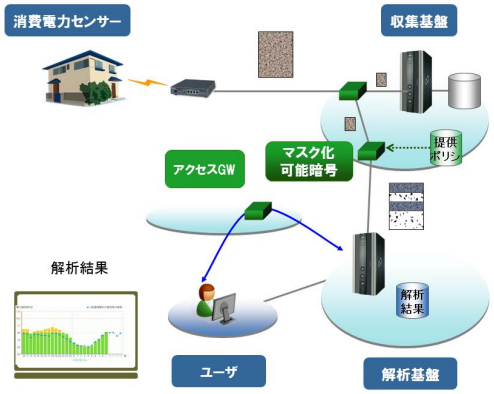

図 5: アクセス GW を用いた構成

3.3.3　仮名 ID の導入

今回の利用シーンでも, アクセス GW はユーザとサービスの間をつなぐ役割を行う. ユーザはアクセス GW にログインし, アクセス GW では, 代理ログイン機能で, 解析基盤にログインする (図 5 参照). その際, ユーザを特定できないように代表アカウントでログインする. ログイン時のみ一時的に利用する仮トークンを解析基盤で発行し, 収集基盤では, その仮トークンと仮名 ID とを関連づかせることで, アクセス GW では仮名 ID を用いなくて済む. また, これにより, 解析基盤では仮トークンを用いて収集基盤に仮名 ID の問い合わせを行い, その仮名 ID に基づく解析データをユーザに提供することで, ユーザの特定をせずに済み, 課題を解決できる.

4　FCDS への適用検討

自動車がセンサーとなって収集した位置, 加速度, 燃料消費などのデータをフローティングカーデータ (FCD), このデータを収集・分析するシステムをフローティングカーデータシステム (FCDS) と呼ぶ. FCDS の例として, 位置情報を用いたカーナビゲーションシステムが挙げられる. 本節では, FCD の例として加速度情報に着目し, 運転分析システムにおけるプライバシー問題を抽出した上で, マスク化可能暗号方式を用いた保護方法を検討する.

4.1　運転分析システム

自動車の加速度情報から, 運転時の急発進・急停車の様子を把握することができる. 急発進・急停車の理由としては, 道路事情によるものと運転手事情によるものがあり, 同じ地点で (運転手に依存せずに) 急発進・急停車が発生する場合には道路事情によるものそうでない場合には運転手事情によるものと推測できる. 従って, 急発進・急停車が道路事情によって発生する場所を地図にマッピングすることで運転危険地域地図を作成するが可能であり, さらにその地図と運転手事情による急発進・急停車情報を照らし合わせることによって, 運転手の運転 (加速度) が適正かを判定する運転分析システムが実現できる. 米国では運転分析結果を加味した自動車保険 (PAYD, Pay As You Drive) が知られており, 日本にも導入の予兆が見られることから, このような安全分析システムの利用も広がっていくと考えられる.

運転分析に際し, 運転手は位置情報を提供するが必須であるが, どのユーザがどこを運転していたかが運転分析者に知られることとなり, プライバシーが問題となる. しかし, 運転分析する上で位置情報の提供は運転危険地域地図を参照するために不可欠であるため, 前節の EMS における検討のように, センサーデータの一部を秘匿化するだけでは不十分である.

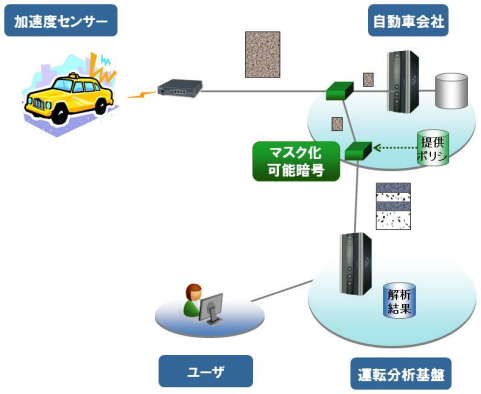

図 6: 運転分析システム

4.2　マスク化可能暗号の適用

運転分析システムにおけるプライバシー保護を実現する上で, 井谷等のマスク化可能暗号方式を適用した構成を図 6 に示す. そこで各自動車は加速度情報を含むフローティングカーデータのセンサーから, 各自動車会社へデータを暗号化して送信することを想定する. 運転分析システムを利用する場合, ユーザは自分の加速度情報や位置情報を運転分析システムに提供することになるが, マスク化可能暗号によって, 自動車会社用に暗号化されたデータを運転分析システム用に暗号化されたデータに再暗号化して提供する.

このままでは運転分析システムがユーザ情報と位置情報を得てしまうため, さらなる対策が必要である. そこで本稿は, 位置情報を丸めること (曖昧化) による対策を提案する. 一般に, 位置情報センサーによって得られる位置情報は詳細であるため, その自動車がいつどこにいたかが明らかになってしまう. 位置情報の一部 (緯度や経度の下位部分) を丸めることで, 「どこ」という情報を「どの辺」に曖昧化すれば, ユーザとの結びつきが困難になると考えられる.

マスク化可能暗号方式においては, 各部分データは暗号化・復号・秘匿化 (マスク化) やその重ね合わせが可能である. 上記のような丸めを実現するためには, 暗号化と位置情報の下位部分へのマスク化を重ねることで, 曖昧化が実現できる.

5 まとめ

本稿では, スマートメータデータの可用性とプライバシー保護の両立を目的として, 井谷等によって提案されたマスク化可能暗号方式をエネルギーマネジメントシステム (EMS) とフローティングカーデータシステム (FCDS) に適用した場合を検討した. これら適用シーンでは, センサーデータは収集基盤に暗号化されて送信され, また外部の解析基盤には, 平文を介することなく不要データの秘匿化と解析基盤向けの暗号鍵に変換されることから, センサーデータのプライバシーが確保できている. さらなる適用シーンの検討は課題である.

参考文献

[1] R. Agrawal, and R. Srikant, "Privacy-Preserving Data Mining", *SIGMOD 2000*, pp.439-450, ACM, 2000.

[2] 富士キメラ総研, "2012 センサデバイス／ソリューションビジネス市場調査総覧", 2012 年 3 月.

[3] C. Gentry, "Fully Homomorphic Encryption Using Ideal Lattices", *STOC 2009*, pp. 169-178, ACM, 2009.

[4] 井谷, 阿比留, 伊豆, 伊藤, 牛田, 小倉, 津田, 武仲, "センサーデータのプライバシ保護つきデータ利活用に適した再暗号化システム", コンピュータセキュリティシンポジウム 2012 (*CSS 2012*), 2012 年 10 月

[5] 伊藤, 牛田, 小櫻, 津田, "ユーザ権限に応じたプライバシー保護データマイニング方式", コンピュータセキュリティシンポジウム 2010 (*CSS 2010*), 3C1-2, 2010 年 10 月.

[6] 平成 23 年度エネルギー管理システム導入促進事業費補助金 (HEMS,BEMS). http://sii.or.jp/hems/, http://sii.or.jp/bems/

[7] NTT, FLET'S ミルエネ. http://flets.com/eco/miruene/

[8] 小倉, 雨宮, 千草, 濱田, 黒川, 阿比留, "安全なデータ・サービス連携システムにおける KVS 技術適用:スケーラビリティ評価", 信学技報 IN 2012-43, pp. 61-66, 2012 年 7 月.

[9] 小倉, 千草, 黒川, 井谷, 阿比留, "他社クラウドを含めた安全なデータ・サービス連携方式の提案", 信学技報 IN 2011-57, pp. 69-74, 2011 年 7 月.

[10] 津田, 松尾, 阿比留, 長谷部, "安全なクラウド連携のためのデータセキュリティ" 雑誌 FUJITSU, Vol. 62, No.5, pp. 531-537, 2011 年 9 月. http://img.jp.fujitsu.com/downloads/jp/jmag/vol62-5/paper10.pdf

[11] 牛田, 伊藤, 片山, 小櫻, 津田, "ゲートウェイによるクラウド間のデータ秘匿集計技術", 2011 年暗号と情報セキュリティシンポジウム (*SCIS 2011*), 3F1-5, 2011 年 1 月