パケット解析におけるプライバシの確保

山崎 拓海[1,a)]　串田 高幸[1]

概要：EU の GDPR をはじめ, 近年, 世界では個人情報に対する扱いがより一層厳しくなっている. その中で IT 業界では, ビッグデータなど, データを用いたビジネスが展開されている. データを集めるうえでいかにして個人情報を守っていくが課題となっている. 本論文ではネットワーク上を流れるパケットをキャプチャし解析するにあたって, 通信者 (送信元, 送信先) を特定できる情報, またメールアドレス, パスワード, WEB サイトの会員情報をマスキングすることで, プライバシーの保護を目指す. LAN での小規模な実験となる. 監視対象のネットワーク内にパケットを取得するマシンを用意する. パケットを取得するソフトウェアとして, tcpdump を用いる. 取得されたパケットはファイルに出力し, 今回保護したい情報が書かれている部分を共通鍵暗号を用いて暗号化し, データ保管用のマシンに渡す. これにより情報に機密性をもたせて, 通信者のプライバシ確保しながらデータを集めることができる.

1. はじめに

近年, 世界中で個人情報の取り扱い方が問題となっている. 2019 年には, Facebook がそれぞれ 2 億 6700 万件以上, 4 億件以上の個人情報を流出させる事態を起こしている. その中で,　EU で個人情報を保護する動きがなされている. 2016 年に可決された ”General Data Protection Regulation” (GDPR) は, これまでの個人情報とされる情報から, 対象となるものが拡大された.[1] これにより, IP アドレス (静的, 動的を問わない), Cookie も個人情報とみなされることになった. また, この法律が適用されるのは EU 加盟国のみにとどまらない. EU 圏の居住者に向けて展開されているサービスであるならば, 会社の所在地にかかわらず GDPR が適用されることになっている. よって, この法律は EU 圏に向けてサービスを展開している世界中の企業に影響を与えることになった. その影響はもちろん日本にも及んでいる. 2018 年に日本の企業向けに行ったアンケートでは, 十分理解していると回答した企業が 10 パーセントであったという話もある. この法律を守らなかった時, 罰則金をはらうことになる. その額は, 全世界での売り上げ高の 4 パーセント または, 2000 万ユーロのいずれか高いほうを最大 1 億ユーロ支払うことになる. 既に Google が約 62 億円の支払いを命じられている.[2]

このように, 個人情報の取り扱いが厳しくなっている. その中で IT 業界では, データを用いたビジネスが展開されている. そのためにはデータを集める必要がある. データを集めることに同意を得ることはもちろんのことだが, 集めるうえでいかにして個人情報を守っていくが課題となっている. しかし, 企業はユーザがもたらす情報を欲している. そこで, ユーザの情報 (宛先, 送信元 IP アドレスおよび MAC アドレス), また WEB サイトの会員情報 (ユーザ ID, パスワード, 通販サイトではクレジットカード情報) を公開鍵暗号方式を用いて特定できない状態にしたうえでデータ保管専用の PC に渡したいえで解析を行うことにより, 通信者のプライバシを守りながら情報を集める方法について提案していくのが本論文である.

以下, 本稿は次のように構成されている. 2 章では GDPR によるビジネスへの影響, プライバシーの保護に関する研究を紹介する. 3 章で, 提案するプライバシ保護の手法を具体的に説明していく. 続く 4 章で, 提案する手法についての課題を挙げ, 評価を行う. 5 章でまとめを述べ本稿のしめとする.

2. 関連研究

GDPR が制定された背景として, デジタル技術の発展に伴い, 個人が特定できる情報への攻撃がかつてないほど増えていることが挙げられている. EU の競争総局の現最高責任者である Margrethe Vestager 氏は”人々は自分の個人情報には価値があるという事実にますます気づいていると思う”と述べている. また Pew Research Center の調査によると, アメリカ合衆国の成人の 93%が, 誰が自分に関する情報を入手できるか管理することが重要であると回答しており, 90%が収集する情報を管理することが重要であると回

[1] 東京工科大学コンピュータサイエンス学部
CDSL, TUT, Hachiouji, Tokyo 101–0062, Japan
[a)] C0117304

答している．この調査や，個人情報への攻撃が増えていることを鑑みると，Margrethe Vestager 氏が述べたことは説得力があり，個人情報を守るための法律が新たに制定されたり改訂される動きが強まるのも頷ける．[3]

3. 提案

それでは，提案する手法について具体的に述べていく．環境としては小規模な LAN を想定しており，有線で接続されているものとする．パケットを取得するためのソフトウェアは”tcpdump”を使用する．

通信にやりとりされるパケットには実際のデータ部分のほかにヘッダといわれる，情報が記載されている部分がある．プライバシを保護するのに隠したい情報が記載されているのは，Ethernet ヘッダと IP ヘッダである．ヘッダは決まった形で形成されているため，何ビット目から目的の情報があるのかは全パケット同じである．ここで，情報をマスキングする方法として公開鍵暗号を用いる．そしてマスキングされた情報はデータ保管用の PC に送信される．このデータ保管用の PC は普段はスタンドアロンの状態にしておき，データを送る時のみ接続を確立する．パケット取得用 PC ではデータの送信した後にパケットデータを削除し，オンライン上にデータが存在しないようにする．以上が提案する手法の一連の流れである．

Version	IHL	Type of Service	Total Length	
Identification			Flags	Fragment Offset
Time to Live		Protocol	Header Checksum	
Source Address				
Destination Address				
Options			Padding	

図 1　IP ヘッダ

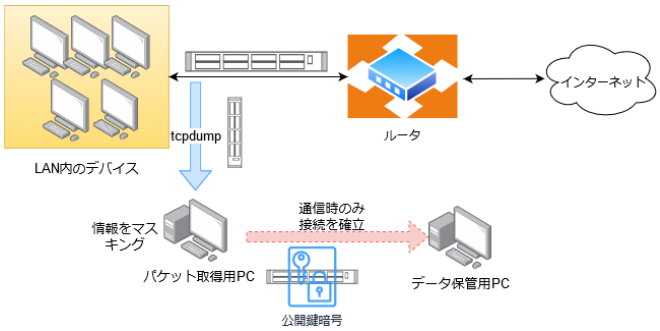

図 2　ソフトウェア構成図

4. 評価，課題

ここでは前セクションで提案手法の評価や課題を挙げていく．

公開鍵を用いるメリットとしてまず，複合される可能性が低いことにある．公開鍵暗号方式での複合には，必ず用いた公開鍵とペアになっている秘密鍵がないと行うことができない．極端な話，秘密鍵を破棄してしまえば暗号化した人ですら複合できなくなるのである．基本的に暗号化した情報を複合することは想定していないが，ネットワークで通信に異常や問題が発生したときに調べる必要がでて場合，複合することで問題の解決の助けになると考えている．課題として，IP アドレスや MAC アドレスを暗号化しただけで通信者のプライバシを保護しているといえるのかという懸念が残っている．その解決策として，ほかのパケットの情報と暗号化する前にスワップしてしまい後に暗号化することによって，より通信者を特定することを困難にしプライバシの保護につながる．

次にデータ保管用の PC についてだが，普段はスタンドアロンとしておきデータの受け渡し時のみ接続確立することで不正アクセスのリスクを限りなくなくすことができる．私が想定している環境としては，データ保管用の PC は LAN 内にあるものになっているが，様々なクラウドサービスが各社より提供されている今日おいてはわざわざ LAN 内にデータ保管用の PC を置いてくのではなく，このようなサービスを利用したほうがよい場合もある．

5. おわりに

本論文でとりあげた GDPR は世界中の企業に影響を与えたものになったが，あくまで EU 圏での法律となっている．このような法律が世界各国で制定される可能性はとても高いとみられる．それでもユーザの嗜好を解析しそれにあったサービスを提供するのが当たり前になっているので，企業としてはデータの解析をとめることはできない．ユーザのデータを多く持っている企業はサイバ攻撃を受ける可能性があり，情報が流出するようなことが起きれば信頼にもかかわることになる．しっかりと自分を守る手立てをもち，個人情報の重要性を十分認知しながら扱っていかなければならない．

参考文献

[1] Colin, T.: What the GDPR means for businesses, Proc. Network Security, pp.5-8 (2016).
[2] Markus, P.: A Method to Compress and Anonymize Packet Traces, (2001).
[3] Jay, B.: How automating data collection can improve cyber-security, Proc. Network Security, pp.11-13, (2017).
[4] Karan K.B., Abhishek, M., Mehdi, B., Chinmay, K., Ashish, K. and Mukesh, S.: Risk-Based Packet Routing for Privacy and Compliance-Preserving SDN, Proc.

IEEE (2017).
[5] Deyan, C. and Hond, Z.: Data Security and Privacy Protection Issues in Cloud Computing, Proc. IEEE (2012).

参考文献

[1] Albrecht, J. P.: How the GDPR Will Change the World, *EDPL*, Vol. 2, pp. 287–289 (2016).
[2] Heller, M.: Google GDPR fine of $57 million sets record, *TechTarget*, (online), available from ⟨https://searchsecurity.techtarget.com/news/252456372/Google-GDPR-fine-of-57-million-sets-record⟩ (2019).
[3] Greengard, S.: Weighing the Impact of GDPR, *Commun. ACM*, Vol. 61, No. 11, p. 16–18 (online), DOI: 10.1145/3276744 (2018).

