1X-02

準同型暗号を用いた秘匿検索のログ解析への応用

井上　翼†　　福田　洋治†　　廣友　雅徳‡　　白石　善明*

近畿大学†　　佐賀大学‡　　神戸大学*

1. はじめに

セキュリティインシデントの原因を調査する手段の一つとしてログ解析があり，これにより，誰がどんな操作を行ったのか，どんなプログラムが動作したのか，システムがどういう状態にあるのか，システムがどんな通信を行ったのか等の当時，起こった事象を時系列に把握できる[1)]．

組織に記録される各種ログは一般に膨大であり，ログ解析には特別な知識，スキルが求められることもあり，ログの管理，解析を外部委託したいという要求がある反面，ログの中には個人のプライバシー情報や組織の秘密情報が含まれることがあるのでプライバシーやセキュリティの懸念が残る[2)]．

ある文章のキーワードが並べられているインデックスから特定のキーワードが存在するかを検索して結果を返す処理で，文章とインデックス，検索キーワードを暗号化したまま扱える準同型暗号を用いた秘匿検索[3)]が存在する．

本研究では，顧客の各種ログを暗号化したかたちでクラウド上に保管し，クラウド上のログを暗号化されたまま監視し，必要に応じて顧客がアラートを受け取ることができる，準同型暗号を用いた秘匿検索[3)]を利用した秘匿ログ解析の手法を検討，提案する．

2. 秘匿ログ解析の手法

準同型暗号を用いた秘匿検索[3)]を利用した秘匿ログ解析の手法を図1に示す．

［事前準備］

(a-1) BGV 方式の完全準同型暗号のためのパラメータλ，ブルームフィルタに登録するレコード数nと偽陽性率pを設定する．

(a-2) λを用いて公開鍵pkと秘密鍵skを生成する．

(a-3) nとpからブルームフィルタのサイズmを，nとmからハッシュ関数の数kを計算する．

［平常時］

(b-1) サイズmのブルームフィルタを7つ生成，それぞれを0で初期化し，それを1つのBF_1とする．その7つは，ログレコードに含まれる 5W1H から When を日付と時間に分けた7つの項目に対応する．

(b-2) ログレコードL_1で，k個のハッシュ関数に，L_1の検索用キーワード$kw1_1$を与えて，各ハッシュ値をBF_1の対応する項目の番地として，そこの要素を1にする．それを，L_1内にある全ての検索用キーワード$kw1_h$（$1 \le h \le h_{max}$，h_{max}：L_1内の検索用キーワード数）で行い，その処理をn個分のログレコードで行う．

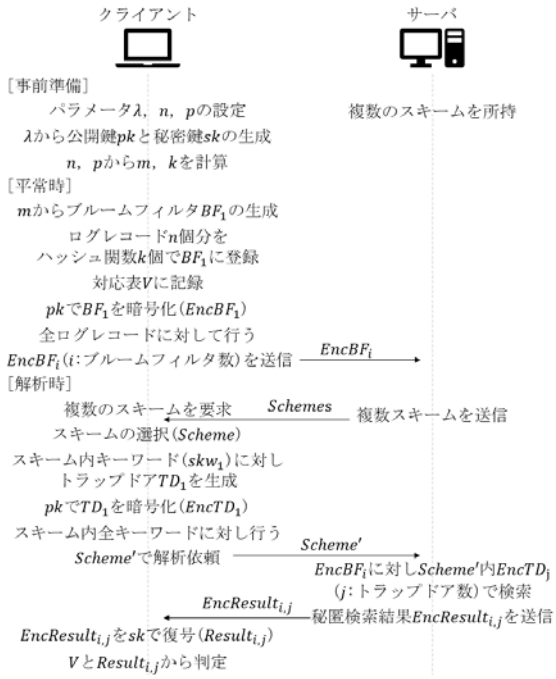

図1　秘匿ログ解析の手法

(b-3) ログレコードn個の登録が完了すると，BF_1が対応しているレコードを対応表Vに記録する．

(b-4) BF_1をpkを用いて暗号化し，$EncBF_1$とする．

(b-5) b-1～b-4 の処理を，登録したいログレコード全てに対して行う．

(b-6) 暗号化された全ブルームフィルタ$EncBF_i$（$1 \le i \le i_{max}$，i_{max}：ブルームフィルタ数）をサーバに送信する．

［解析時］

(c-1) クライアントはサーバが所持している複数のスキームを取得し，スキーム$Scheme$を選択する．

(c-2) スキームには，検索したい項目とキーワード，一致具合，AND や OR の情報が載せられている．その内のキーワードskw_1に対してサイズmのトラップドアTD_1を生成，0で初期化する．k個のハッシュ関数にskw_1を与えて，各ハッシュ値をTD_1の番地として，そこの要素を1にする．

(c-3) TD_1をpkを用いて暗号化し，$EncTD_1$とする．

(c-4) c-2～c-3 の処理を，スキーム内にある全てのキーワードに対して行う．

(c-5) キーワードが加工されたスキームを$Scheme'$

An Application of Secret Search using Homomorphic Encryption to Log analysis

† Tsubasa INOUE, Youji FUKUTA, Kindai University

‡ Masanori HIROTOMO, Saga University

* Yoshiaki SHIRAISHI, Kobe University

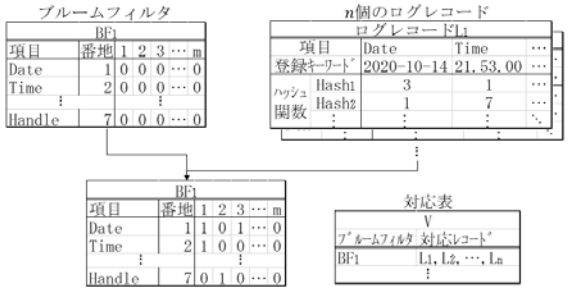

図2　ブルームフィルタ作成の様子と対応表

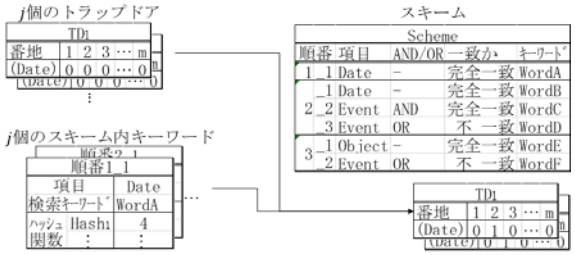

図3　トラップドア作成の様子とスキーム

とし，それでサーバに解析依頼を行う．

(c-6)保管されている全ての暗号化ブルームフィルタ$EncBF_i$に対して，$Scheme'$内の各暗号化トラップドア$EncTD_j$（$1 \le j \le j_{max}$，j_{max}：$Scheme'$内の暗号化トラップドア数）で順に検索をする．検索のアルゴリズムは，$EncTD_j$に紐付いている項目と同じ項目を$EncBF_i$から選択し，その各要素をNOT演算する．それと，$EncTD_j$を同じ番地同士で乗算をする．最後に，各番地の値が全て0ならば一致となる．

(c-7)検索結果は暗号化されている($EncResult_{i,j}$)ので，サーバは検索の結果を知れない．それを，クライアントに送信する．

(c-8)返された依頼結果をskで復号する($Result_{i,j}$)．

(c-9)記録をしていた対応表Vと，復号された検索結果$Result_{i,j}$から，$Scheme$に関する攻撃がされているかを判定する．判定結果が不十分であれば，c-1から別のスキームを選択して再度行う．

3.　データサイズの考察と計算時間の実験

平常時の暗号化前のブルームフィルタのサイズをmとおく．ブルームフィルタ1つに登録するログレコード数nと偽陽性率pに左右され，$m = -(n \ln p)/(\ln 2)^2$となる．ログレコード内のキーワードを7項目に分け，項目ごとにブルームフィルタを用意すると考えると，$m * 7$ が必要なブルームフィルタのサイズとなる．これに，暗号化時のデータサイズ増加率eを掛けた物が，$\lfloor (l-1)/n \rfloor + 1$個（$l$:ログレコード数）生成される．$n$個のログレコードに対して生成される暗号化ブルームフィルタのデータサイズは$(m * 7) * e * (\lfloor (l-1)/n \rfloor + 1)$と書ける．1レコード内の登録キーワードの平均総文字数をxとおくと$(m * 7)/(x * n)$のデータサイズ軽減となる．

解析時の1つのトラップドアのサイズはmである．ログレコードの，ある1項目を検索するためのトラップドアのサイズはmとなる．1つのスキームの検索キーワード数がt個ある場合，必要なトラップドアのサイズは$m * t$となり，暗号化をするとeが掛けられる．それに項目情報や，ANDやOR条件，一致具合のデータサイズ cが加わり，$(m * t) * e + c$が解析依頼時のスキームのデータサイズとなる．

サーバは解析時，保管されている暗号化ブルームフィルタ$\lfloor (l-1)/n \rfloor + 1$個に対し，スキーム内トラップドア数$t$を掛け，ブルームフィルタサイズ$m$を掛け，暗号化時の$e$を掛けたサイズ$(\lfloor (l-1)/n \rfloor + 1) * t * m * e$のデータをクライアントに返す．

準同型暗号ライブラリ HElib[4]，C++言語を用いて提案手法の各手続きの暗号化の処理をプログラムで記述し，OS:Ubuntu18.04LTS，CPU:Intel Core i5-5250U 1.6GHz，Memory:3.9GBのPC上で動作させて，実行時間を計測する実験を行った．ブルームフィルタやトラップドアの各要素は2値で，それに適した準同型暗号のパラメータを与え，$n = 100, p = 0.001$でレコード数を5000で動作させた．準同型暗号の準備部分が 262[ms]，ブルームフィルタの作成・暗号化が 73647[ms]となった．またトラップドアを1つだけ作成・暗号化すると202[ms]となった．既に作成されたブルームフィルタとトラップドアを用いると，検索処理は26138 [ms]となる．

4.　おわりに

顧客の各種ログを暗号化したかたちでクラウド上に保管し，クラウド上のログを暗号化されたまま監視し，必要に応じて顧客がアラートを受け取ることができる，準同型暗号を用いた秘匿検索[3]を利用した秘匿ログ解析の手法を提案した．

準同型暗号を用いた秘匿検索[3]を利用した秘匿ログ解析の手法でログレコード以外に作成されるブルームフィルタ，トラップドアのデータサイズについて考察し，主要な計算を占める準同型暗号の処理時間を，PC 上で試作プログラムを動作させて，計測したところ，それほど大きな処理時間とはならず，実現可能な余地があることを確認した．

参考文献

1) 満永 拓邦:高度サイバー攻撃への対処におけるログの活用と分析方法，JPCERT/CC，available at https://www.jpcert.or.jp/research/APT-loganalysis_Report_20161019.pdf，(参照 2020/10/18).
2) 石田 茂:プライバシーに配慮したアプリケーションログ出力の設計，情報処理学会研究報告，Vol.2015-EIP-68 No.10，2015.
3) 柴山 綸太郎，土井 洋:公開鍵検索可能暗号に適したブルームフィルタの検討，第18回情報科学技術フォーラム，第4冊分，pp.127-132，2019.
4) Github，homenc/HElib，available at https://github.com/homenc/HElib（参照 2020/10/03).