Giải pháp lưu vết và thu hồi thiết bị thu bất hợp pháp : Luận văn ThS / Nguyễn Thị Ngọc Mai ; Nghd. : PGS.TS. Trịnh Nhật Tiến . - H. : ĐHCN, 2007 . - 119 tr. + CD-ROM

MỞ ĐẦU

Hiện nay vấn đề bảo vệ bản quyền đang là vấn đề nhức nhối của Việt Ị am cũng như trên thế giới. Vấn đề bảo vệ bản quyền với các tác phẩm công nghệ số là vấn đề mà luật bản quyền phải đương đầu do: công nghệ số tạo khả năng cho việc truyền phát và sử dụng tất cả các đối tượng bảo hộ của bản quyền và quyền kế cận dưới dạng số dễ dàng. Quy trình số hóa cho phép biến đổi các tác phẩm này thành dạng nhị phân, khiến cho chúng dễ dàng được truyền qua mạng Internet và sau đó được phân phối, sao chép, và cất giữ một cách hoàn hảo dưới dạng số.

Các thách thức trên diễn ra đối với ngành công nghiệp bản quyền, khi mà số tiền thu được từ bản quyền trong nền kinh tế quốc dân đang đạt tới mức khó dự đoán trước. Giá trị kinh tế của riêng ngành công nghiệp bản quyền tại Mỹ ước tính đạt 91,2 tỷ đô la Mỹ (theo thông tin từ Liên minh Sở hữu trí tuệ thế giới (IIPA)) chiếm tới 5,24% tổng sản phẩm quốc nội của Mỹ, tăng nhanh gấp 2 lần phần còn lại của nền kinh tế.

Giá trị của ngành công nghiệp bản quyền chiếm 6% giá trị tăng thêm của nền kinh tế Uruguay vào năm 1997, chiếm 6,7% giá trị tăng thêm của nền kinh tế Bra-xin vào năm 1998, thu hút 1.3 triệu việc làm tại quốc gia này.

Tại Việt ả am, tỷ lệ vi phạm bản quyền đối với các tác phẩm nghe nhìn rất nghiêm trọng gây tổn thất cho nền kinh tế, gây khó khăn trong quá trình hội nhập với thế giới. Chính vì vậy, việc tìm kiếm các giải pháp kỹ thuật và luật pháp để bảo hộ bản quyền khỏi sự sao chép "số" bất hợp pháp là vô cùng cấp thiết.

ả hiệm vụ của luận văn này là trình bày các giải pháp kỹ thuật về lưu vết và thu hồi các thiết bị thu bất hợp pháp, nhằm bảo vệ bản quyền, bảo vệ nội dung của các tác phẩm được truyền phát qua các kênh quảng bá (broadcast channels).

LUẬN VĂN GỒM CÁC NỘI DUNG SAU:

MỞ ĐẦU

Chương 1: CÁC KHÁI NIỆM CƠ BẢN

Giới thiệu các khái niệm cơ bản sử dụng trong luận văn. Chương này cũng nêu lên các thành phần cơ bản của một hệ thống phát dữ liệu quảng bá.

Chương 2: GIẢI PHÁP LƯU VẾT THIẾT BỊ THU LÀM RÒ RỈ KHÓA

Chương này trình bày giải pháp lưu vết thiết bị thu làm rò rỉ khóa bí mật, sử dụng phương pháp phân hoạch tập TBTDL thành các tập con.

Chương 3: GIẢI PHÁP THU HỒI THIẾT BỊ THU BẤT HỢP PHÁP

Trình bày giải pháp thu hồi thiết bị thu bất hợp pháp sử dụng khung phủ tập con (Subset Cover Framework).

Sau đó sẽ trình bày các giải pháp áp dụng khung phủ tập con là: Cây nhị phân đầy đủ (Complete Subtree), Hiệu hai tập con (Subset Difference).

Chương 4: ĐỘ AN TOÀN CỦA KHUNG PHỦ TẬP CON

Chương 5: MỘT SỐ ỨNG DỤNG

KẾT LUẬN

Chương 1: CÁC KHÁI NIỆM CƠ BẢN

1.1 MỘT SỐ KHÁI NIỆM

- ***Trung tâm quảng bá (Center, Broadcast Center), nhà cung cấp dữ liệu (NCCDL – Data Provider)*** [6]: Trung tâm có các kênh phát thông tin quảng bá tới các thiết bị thu dữ liệu.
- ***Thiết bị thu dữ liệu (TBTDL - User)*** [3]: thu dữ liệu phát ra từ ả CCDL và dùng các khóa bí mật của nó để giải mã dữ liệu thu được.
- ***Thông điệp hay bản tin (Message)*** [6]: là thông tin hoặc đoạn thông tin được ả CCDL gửi đến TBTDL qua các kênh quảng bá.
- ***Khóa thời gian tồn tại ít (short-lived key - session key)*** [3]: là khóa được duy trì trong một phiên truyền dữ liệu gọi tắt là **khóa phiên**.
- ***Khóa thời gian tồn tại dài (long-lived key)*** [3]: là khóa tồn tại trong thời gian dài của hệ thống, gọi tắt là **khóa thời gian dài** hay **khóa "dài"**.
- ***Bộ khóa nhái*** [2]: Là bộ khóa mà kẻ gian đã thu được từ tập khóa của một số TBTDL (bằng phương pháp nào đó, ví dụ thám khóa).
- ***Thiết bị thu bất hợp pháp (Traitor)*** [2]: là TBTDL làm rò rỉ khóa hoặc TBTDL sử dụng bộ khóa nhái để giải mã bản tin nhận được từ ả CCDL.
- ***Truyền tin quảng bá*** (***Broadcast, Transmistion)*** [6]: quá trình ả CCDL phát định kỳ thông điệp đã mã hóa tới TBTDL.
- ***Giải pháp lưu vết TBTDL làm rò rỉ khóa (Tracing Traitor)*** [2]: xác định định danh TBTDL làm rò rỉ khóa.
- ***Giải pháp thu hồi TBTDL bất hợp pháp (Revocation Traitor)*** [3]: là giải pháp phân hoạch các TBTDL hợp pháp thành các tập con, dựa vào đó ả CCDL mã hóa thông điệp, để TBTDL bất hợp pháp không giải mã chính xác thông điệp ả CCDL phát quảng bá.

Các ký hiệu dùng trong luận văn:

- ẩ : tập tất cả các TBTDL do ẩ CCDL quản lý, |ẩ |= n.
- $u_1,...,u_n$: ký hiệu các TBTDL thuộc ẩ .
- R: tập các TBTDL làm rò rỉ khóa, |R|= r.
- P: tập các TBTDL hợp pháp, P=ẩ - R.
- K: khóa phiên (session key hay short-lived key).
- L: khóa "dài" (long-lived key).
- M: thông điệp hay bản tin.
- C_M: bản mã của thông điệp M.
- tM: bản tin thử nghiệm (test message).
- L_{u_i} : tập các khóa "dài" của TBTDL u_i , i=1, 2,…, n.
- $|L_{u_i}|$: số lượng các khóa "dài" của TBTDL u_i .
- S_i : tập các TBTDL dùng chung một khóa "dài" L_i .
- $S_{i,j} = S_i - S_j$: chứa các TBTDL thuộc phần bù của tập S_i so với tập S_j .

  Các TBTDL trong tập $S_{i,j}$ dùng chung khóa "dài" $L_{i,j}$.

Định nghĩa phủ [6]

Cho một họ các tập con khác rỗng
$S = \{S_1, S_2,..., S_w\}, S_j \subseteq ẩ , j = 1,..., w$.

Cho tập khác rỗng $P \subseteq ẩ$; phủ của tập P là tập $S_{i_1}, S_{i_2},..S_{i_t}$, $\{i_1, i_2,...i_t\} \subseteq \{1,..., w\}$ và thỏa mãn điều kiện:

$$P = \bigcup_{j=1}^{t} S_{i_j}, \quad S_{i_j} \cap S_{i_k} = \phi, \quad \forall i_j \neq i_k$$

Kích thước của một phủ là số lượng các tập con tạo nên phủ đó. Ví dụ ở đây, kích thước của phủ P là t.

1.2 KHUNG PHỦ TẬP CON

Phần này giới thiệu **khung phủ tập con (Subset Cover Framework - SCF)** được sử dụng trong giải thuật lưu vết và giải thuật thu hồi TBTDL bất hợp pháp [3].

Trong SCF, có giải thuật xác định các tập con $S_1, S_2, ..., S_w \subset$ ẩ , $\bigcup_{i=1}^{w} S_i =$ ẩ . Mỗi tập S_i có khóa "dài" L_i.

Tập P phải được phân hoạch thành các tập con rời rạc $S_{i_1}, S_{i_2}, ..., S_{i_m}$ sao cho: $P = \bigcup_{j=1}^{m} S_{i_j}$.

Các khóa "dài" tương ứng với các tập $S_{i_1}, S_{i_2}, ..., S_{i_m}$ là $L_{i_1}, L_{i_2}, ..., L_{i_m}$.

Lưu ý: Các TBTDL $u \in S_{i_j}$ sử dụng chung khóa "dài" L_{i_j}, j=1, 2,…, m

Mỗi $u \in S_i$ đều tính được L_i từ tập khóa L_u của mình.

SCF sử dụng hai giải thuật mã hóa E và F:

- Giải thuật $E : \{0,1\}^* \rightarrow \{0,1\}^*$, mã hóa khóa phiên K, lần lượt với từng khóa "dài" $L_{i_1}, L_{i_2}, ..., L_{i_m}$, nhận được các bản mã: $E(K, L_{i_1})$, $E(K, L_{i_2})$, …, $E(K, L_{i_m})$.
- Giải thuật $F : \{0,1\}^* \rightarrow \{0,1\}^*$, mã hóa thông điệp M sử dụng khóa phiên K, nhận được bản mã: $F_K(M)$.

SCF gồm ba phần [3]:

1) Khởi tạo

- ẩ CCDL có giải thuật xác định các tập con $S_1, S_2, ..., S_w \subset$ ẩ , $\bigcup_{i=1}^{w} S_i =$ ẩ . Mỗi tập S_i có khóa "dài" L_i . L_i có thể là:
  - + Hoặc là số độc lập, ngẫu nhiên: $L_i = G$.
  - + Hoặc là hàm của số độc lập, ngẫu nhiên: $L_i = f(G)$.

  Trong đó G là số được sinh từ bộ sinh số ngẫu nhiên.
- Mỗi TBTDL $u \in S_i$ được ẩ CCDL cấp một tập các khóa "dài" L_u. L_u có thể là:
  - + Hoặc là $L_u = \{L_i\}$, $1 \le i \le w$.
  - + Hoặc là $L_u = \{f(L_i)\}$, $1 \le i \le w$, f : ánh xạ 1-1.

***2) Mã hóa*:** Thực hiện tại ẩ CCDL

- ẩ CCDL chọn khóa phiên K ngẫu nhiên.
- ẩ CCDL phân hoạch P thành các tập con rời rạc $S_{i_1}, S_{i_2}, ..., S_{i_m}$.

  $L_{i_1}, L_{i_2}, ..., L_{i_m}$ là các khóa "dài" tương ứng với các tập con đó.
- ẩ CCDL mã hóa khóa phiên K, một lần với từng khóa $L_{i_1}, L_{i_2}, ..., L_{i_m}$ và phát quảng bá bản mã:

$$< [i_1, i_2, ..., i_m, E(K, L_{i_1}), E(K, L_{i_2}), ..., E(K, L_{i_m})], F_K(M) >$$

Phần trong dấu [ ] gọi là **phần đầu** (header), $F_K(M)$ gọi là **phần thân** (body).

***3) Giải mã*:** Thực hiện tại TBTDL u.

- TBTDL u nhận được bản mã:
  $<[i_1, i_2, ..., i_m, C_{i_1}, C_{i_2}, .., C_{i_m}], M'>$
- u tìm i_j sao cho $u \in S_{i_j}$, trong đó tập các TBTDL hợp pháp là $P = \bigcup_{j=1}^{m} S_{i_j}$. ẩ ếu $u \in R$ thì không tìm được i_j.
- u tính khóa L_{i_j} từ tập khóa L_u.
- Giải mã $D_{L_{i_j}}(C_{i_j})$ để thu được K, $\forall j = 1,2,...,m$.
- Giải mã $D_K(M')$ để thu được M.

1.3 GIẢI PHÁP LƯU VẾT TBTDL LÀM RÒ RỈ KHÓA

a. Bài toán lưu vết

ẩ CCDL truyền thông điệp M qua các kênh quảng bá tới tập ẩ gồm các TBTDL (|ẩ |=n). Mỗi TBTDL u_i có một tập khóa "dài" (bí mật) L_{u_i} (i=1, 2,…, n). Trong tập ẩ có tập R các TBTDL làm rò rỉ khóa bí mật (rò rỉ toàn bộ hoặc một vài khóa trong tập khóa "dài" L_{u_i}) và tập P các TBTDL hợp pháp.

P, R thỏa mãn: $P \cup R = ẩ$, $P \cap R = \phi$.

Yêu cầu ẩ CCDL xác định được định danh các TBTDL thuộc R và phân hoạch P thành các tập con chứa các TBTDL hợp pháp.

b. Giải pháp lưu vết

Có nhiều giải pháp để lưu vết TBTDL làm rò rỉ khóa bí mật. Giải pháp lưu vết TBTDL làm rò rỉ khóa trong luận văn sử dụng khung phủ tập con (SCF - xem 1.2) [3]. Giải pháp đó gồm các phần: ***Khởi tạo, Mã hóa, Giải mã và Thuật toán lưu vết TBTDL làm rò rỉ khóa bí mật.***

- ***Thuật toán lưu vết TBTDL làm rò rỉ khóa bí mật:***

Xác định định danh của TBTDL làm rò rỉ khóa bí mật dựa trên sự phân hoạch tập TBTDL thành các tập con.

Giải thuật này sẽ được trình bày trong chương 2.

1.4 GIẢI PHÁP THU HỒI THIẾT BỊ THU BẤT HỢP PHÁP

a. Bài toán thu hồi

ẩ CCDL truyền thông điệp M qua các kênh quảng bá tới tập ẩ gồm các TBTDL (|ẩ |=n). Mỗi TBTDL u_i có một tập khóa "dài" L_{u_i} (i=1, 2,…, n). ẩ CCDL đã biết tập R các TBTDL làm rò rỉ khóa và tập P các TBTDL hợp pháp.

P, R thỏa mãn: $P \cup R = ẩ$, $P \cap R = \phi$.

Yêu cầu:

Mọi TBTDL thuộc P giải mã chính xác thông điệp M.

Mọi TBTDL thuộc R giải mã chỉ thu được $M^R \neq M$.

Mọi TBTDL sử dụng bộ khóa nhái chỉ thu được M' ≠ M.

b. Giải pháp thu hồi

Có nhiều giải pháp để thu hồi TBTDL bất hợp pháp. Trong luận văn trình bày giải pháp thu hồi TBTDL bất hợp pháp sử dụng khung phủ tập con (SCF - xem 1.2) [3]. Giải pháp đó gồm các phần: ***Khởi tạo, Mã hóa, Giải mã***.

❖ **Hiệu năng của giải thuật thu hồi TBTDL bất hợp pháp**

Hiệu năng của giải thuật thu hồi thể hiện trên ba tham số sau [3]:

- Số lượng khóa "dài" cần lưu trữ tối đa tại mỗi TBTDL u là: $|L_u|$.
- Độ dài phần đầu bản mã (header) trong thông điệp đã mã hóa.
- Thời gian giải mã bản mã thông điệp tại TBTDL.

1.5 MỘT SỐ CÔNG CỤ

1.5.1 Đồ thị

Đồ thị vô hướng G (hình 1.1 a) là một cặp (V, E) [6], trong đó:

+ V là tập các đỉnh hoặc nút, ký hiệu V(G).
+ E là tập các cạnh nối hai đỉnh không phân biệt thứ tự, ký hiệu E(G).

Đồ thị có hướng G (hình 1.1 b) là một cặp (V, E) [6], trong đó:

+ V là tập các đỉnh hoặc nút, ký hiệu V(G).
+ E là tập các cạnh nối hai đỉnh có phân biệt thứ tự. Một cạnh bắt đầu ở một nút và kết thúc ở một nút khác, ký hiệu E(G).

Hình 1. 1: Đồ thị G

Đồ thị liên thông là đồ thị có ít nhất một đường đi giữa 2 nút bất kỳ. **Đồ thị đơn** là đồ thị mà giữa hai đỉnh chỉ có tối đa một cạnh. **Đồ thị có chu trình** (cyclic) là đồ thị tồn tại một đường đi theo dạng: v_1, e_1, v_2, e_2,…, e_i, v_1 (nút kết thúc trùng nút bắt đầu đường đi).

1.5.2 Cây nhị phân

a. Khái niệm cây

Cây là đồ thị đơn, vô hướng, liên thông và không có chu trình [6].

b. Khái niệm cây nhị phân

Cây nhị phân (hình 1.2) là cây có hai dạng nút:

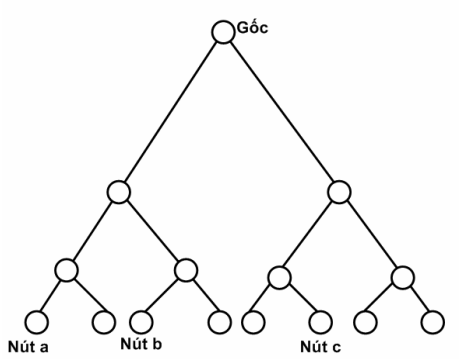
- ả út ngoài: nút lá, không có con.
- ả út trong: có chính xác hai con là con trái và con phải.

Cây nhị phân đầy đủ (hình 1.3) là cây nhị phân, trong đó tất cả các lá có cùng khoảng cách tới gốc.

Số lượng các lá trong cây nhị phân đầy đủ có chiều cao k là $h = 2^k$.

Cha chung thấp nhất của hai nút (kể cả lá) a, b (hình 1.4) là nút giao nhau giữa đường đi từ a tới gốc và từ b tới gốc.

Cây con của cây T là đồ thị con của T và thỏa mãn các tính chất của một cây.



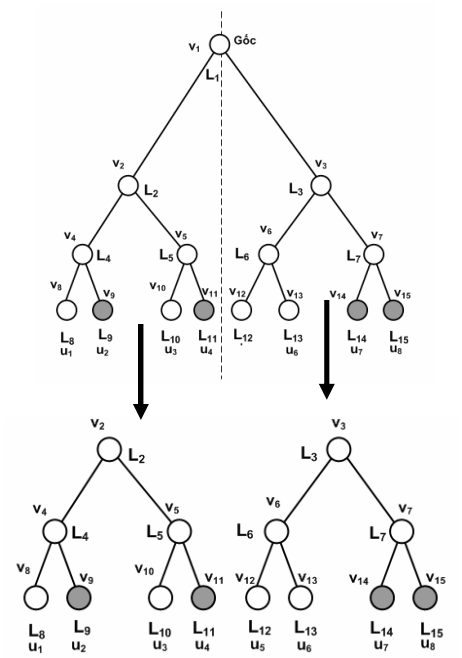
Hình 1.3: Cây nhị phân đầy đủ

c. Tính chất cây nhị phân

1) Cây nhị phân có r lá thì có chiều cao ít nhất là $\lceil \log_2(r) \rceil$

2) Thuộc tính rẽ nhánh

Một cây nhị phân luôn có thể chia thành hai nhánh trái và phải (hình 1.5). Thuộc tính này được gọi là thuộc tính rẽ nhánh của cây nhị phân [3].

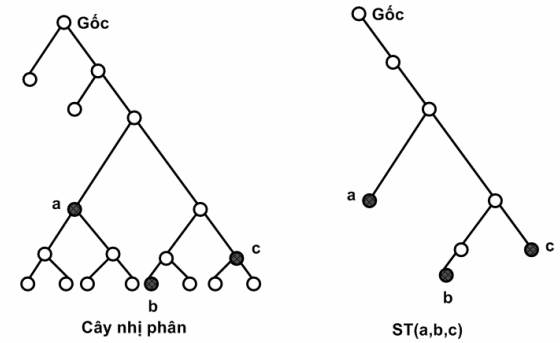


Hình 1.5: Minh họa thuộc tính rẽ nhánh

1.5.3 Cây Steiner

Cho tập R các nút thuộc cây, đồ thị nối gốc và các nút thuộc R được gọi là cây Steiner của R, ký hiệu ST(R).

Ví dụ với tập R = {a, b, c}, ta có cây Steiner ST(a, b, c) như hình 1.6.



Hình 1.6: Cây Steiner của 3 nút (a,b,c)

Chương 2: GIẢI PHÁP LƯU VẾT TBTDL LÀM RÒ RỈ KHÓA

2.1 KHÁI NIỆM LƯU VẾT TBTDL LÀM RÒ RỈ KHÓA

Giả sử kẻ gian, lấy được bộ khóa hoặc một phần bộ khóa từ t TBTDL hợp pháp, họ xây dựng nên bộ khóa giải mã bất hợp pháp và nhân bản thành nhiều bộ khóa nhái[2]. ẩ những TBTDL dùng bộ khóa nhái này và kể cả t TBTDL làm rò rỉ khóa gọi là những **TBTDL bất hợp pháp**.

Để đối phó với tình hình trên, ẩ CCDL tạo ra các TBTDL để làm thí nghiệm "**tại gia**" với các bộ khóa nhái.

ẩ CCDL thu mua được các bộ khóa nhái, họ cho các TBTDL thí nghiệm (**TBTDL_TN**) dùng các bộ khóa nhái này.

ẩ CCDL dùng giải thuật SCF (xem 1.2) phát bản tin thử nghiệm tM, theo dõi những phản ứng của các TBTDL thí nghiệm này, để truy tìm TBTDL đã làm rò rỉ khóa "dài" ra bên ngoài, phân hoạch tập P các TBTDL hợp pháp thành các tập con, để khi ẩ CCDL phát bản tin chính thức, thì các TBTDL bất hợp pháp sẽ không giải mã được chính xác bản tin [2].

***Chú ý*:** ẩ CCDL đã có toàn bộ cấu trúc cây nhị phân T mô tả tập ẩ các TBTDL với giả thiết $n = |ẩ| = 2^k$, các lá tương ứng với các TBTDL. Các nút kể cả lá trong cây được gán tên gọi $v_1, v_2, ..., v_{2n-1}$. Các nút $v_1, v_2, ..., v_{2n-1}$ được gán nhãn $L_1, L_2, ..., L_{2n-1}$.

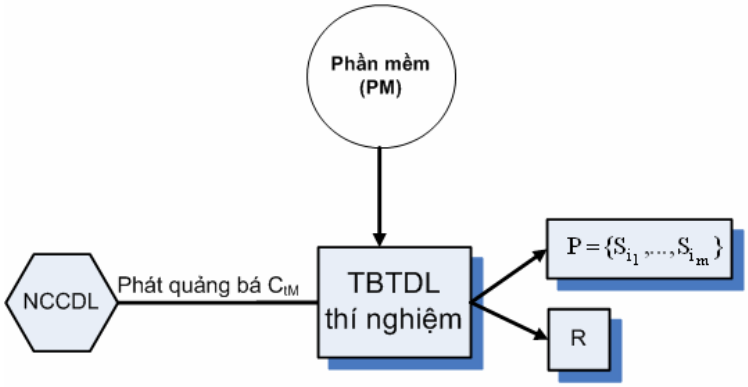
SCF duy trì các tập $S_1, S_2, ..., S_w \subset ẩ$, $\bigcup_{i=1}^{w} S_i = ẩ$. Trong đó S_i là tập các lá của cây nhị phân con gốc v_i, S_i có khóa "dài" L_i tương ứng với nhãn tại nút v_i.

Phương pháp lưu vết của NCCDL đối với một TBTDL_TN:

ẩ CCDL phát thử nghiệm thông điệp tM, PM (Phần mềm) quan sát xác suất giải mã của TBTDL_Tẩ để xác định tập R chứa các TBTDL làm rò rỉ khóa "dài" và phân hoạch tập P các TBTDL hợp pháp thành $P = \{S_{i_1}, S_{i_2}, ..., S_{i_m}\}$. Tìm được P, R thì PM lưu P, R vào cơ sở dữ liệu (**CSDL**) của ẩ CCDL để phục vụ cho lần lưu vết tiếp theo [3].

Giải pháp này giả thiết TBTDL_Tả không thể phát hiện được nó đang bị thử nghiệm, tức là nó không thể tự động tắt máy khi đang thu dữ liệu thử nghiệm từ ả CCDL [3].

Trên thực tế có nhiều giải pháp lưu vết TBTDL làm rò rỉ khóa "dài", luận văn này trình bày giải pháp lưu vết TBTDL làm rò rỉ khóa "dài" dựa trên phương pháp tìm kiếm nhị phân và dùng giải thuật SCF (xem 1.2) để phát bản tin thử nghiệm tM tới TBTDL_Tả . Kết quả được P gồm các tập con chứa TBTDL hợp pháp, $P = \{S_{i_1}, S_{i_2}, ..., S_{i_m}\}$, và tập R các TBTDL làm rò rỉ khóa "dài". Giải pháp này được gọi là giải pháp lưu vết sử dụng tập con, được trình bày trong phần 2.2.



Hình 2.2: Mô hình lưu vết TBTDL làm rò rỉ khóa "dài"

2.2 GIẢI THUẬT LƯU VẾT SỬ DỤNG TẬP CON

2.2.1 Giải thuật lưu vết sử dụng tập con (Subset Tracing)

Ý tưởng của giải thuật là:

Tìm TBTDL làm rò rỉ khóa "dài" bằng cách phân hoạch tập các TBTDL thành tập P và R. Trong đó $P = \{S_{i_1}, S_{i_2}, ..., S_{i_m}\}$ gồm các tập con chứa TBTDL hợp pháp, R là tập các TBTDL làm rò rỉ khóa "dài".

Đầu tiên, giải thuật lưu vết được thực hiện với $P=\{S_1\}$ (S_1 là tập các TBTDL), $R=\varnothing$. Sau khi thực hiện k lần, sẽ được phân hoạch $P = \{S_{i_1}, S_{i_2}, ..., S_{i_m}\}$ và tập R các TBTDL làm rò rỉ khóa "dài". Tập P và R được lưu vào CSDL của ả CCDL.

Tại lần k+1, ả CCDL thu mua được bộ khóa nhái, đưa vào dùng trong TBTDL_Tả , PM sẽ sử dụng hàm Tim_j (mục 2.2.2) để tìm tập con chứa TBTDL làm rò rỉ khóa "dài". Kết hợp với tập $P = \{S_{i_1}, S_{i_2}, ..., S_{i_m}\}$, R trước đó trong CSDL của ả CCDL để xác định P, R mới.

ả ếu TBTDL_Tả giải mã bản tin thử nghiệm tM với xác suất p< 1 thì kết thúc, P và R giữ nguyên, ả CCDL yên tâm vì bộ khóa nhái không có tác dụng.

ả gược lại, tức là TBTDL_Tả giải mã tM với xác suất p=1. Điều đó chứng tỏ bộ khóa nhái có chìa khóa "dài" L_{i_j}, nhờ nó mà TBTDL_Tả đã giải mã được khóa phiên K. Khóa "dài" L_{i_j}, chắc chắn phải do TBTDL nào đó trong tập S_{i_j}, đã làm rò rỉ ra ngoài. Vì vậy PM thực hiện thủ tục Tim_j để tìm chỉ số j sao cho S_{i_j} có chứa TBTDL làm rò rỉ khóa "dài" L_{i_j}.

ả ếu S_{i_j} chỉ chứa một TBTDL thì chắc chắn TBTDL này đã làm rò rỉ khóa "dài" L_{i_j}. Khi đó ta có $R=R \cup S_{i_j}$, và loại bỏ S_{i_j} khỏi tập P. ả gược lại, tức là $|S_{i_j}|>1$, PM chia S_{i_j} - thành hai tập bằng nhau, bổ sung hai tập này vào P, loại bỏ S_{i_j} khỏi tập P.

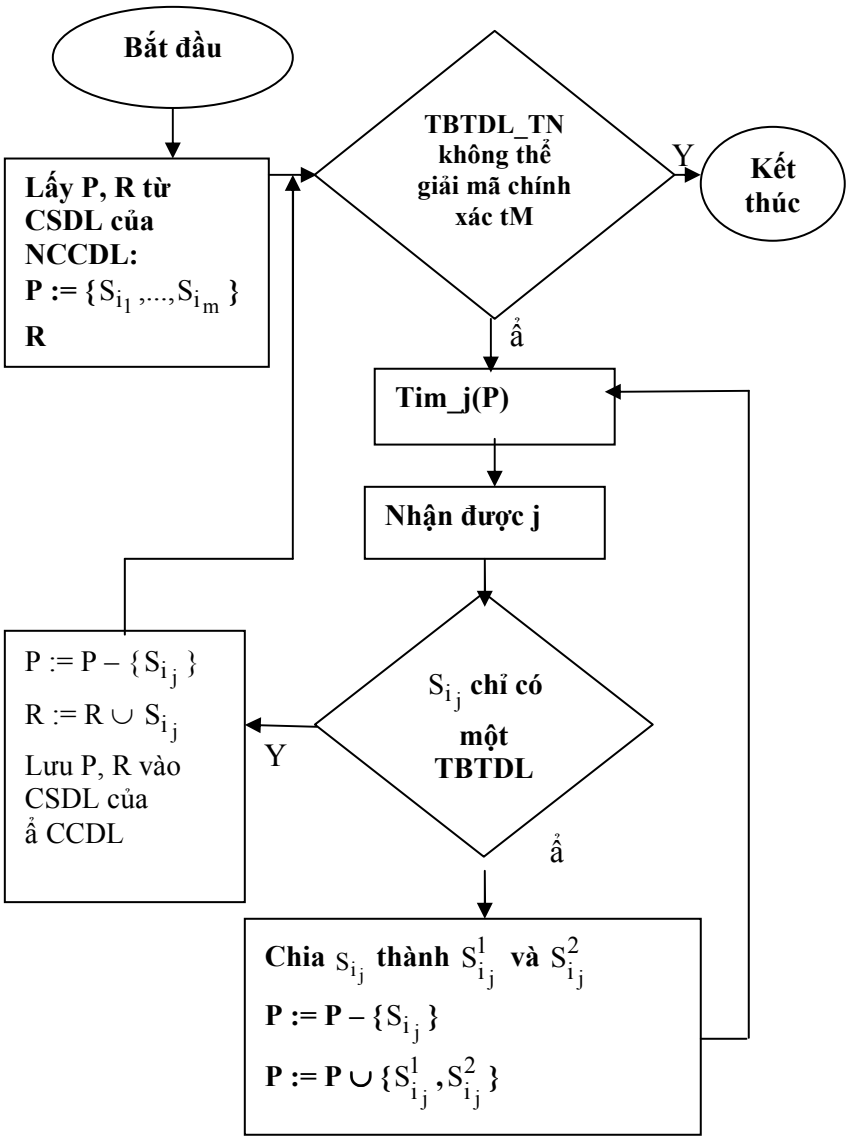
Tiếp tục thực hiện giải thuật lưu vết với phân hoạch P mới, cho đến khi TBTDL_Tả giải mã tM với xác suất p< 1, thì lưu P, R mới vào CSDL và kết thúc lưu vết đối với TBTDL_Tả này.

Phương pháp lưu vết trong luận văn chỉ có tính chính xác tương đối: có thể xác định được định danh của TBTDL làm rò rỉ toàn bộ bộ khóa "dài", hoặc chỉ có thể xác định được khóa "dài" nào bị rò rỉ (vì một số TBTDL chỉ làm rò rỉ một phần của bộ khóa). Dựa trên đó PM phân hoạch P thành các tập con chứa các TBTDL hợp pháp.

Mục đích chính của phương pháp lưu vết này là lưu lại định danh của những TBTDL làm rò rỉ khóa và trừng trị những TBTDL bất hợp pháp. ả ghĩa là danh sách các TBTDL làm rò rỉ khóa sẽ được lưu trong CSDL và P được phân hoạch thành các

tập con chứa các TBTDL hợp pháp để khi ả CCDL phát thông điệp M, những TBTDL bất hợp pháp sẽ chỉ giải mã được M’ ≠ M.

Lưu đồ giải thuật lưu vết [4]



2.2.2 Hàm tìm tập con chứa TBTDL làm rò rỉ khóa

Hàm Tim_j tìm tập con chứa TBTDL làm rò rỉ khóa, được mô tả giống như phương pháp tìm kiếm nhị phân, để tìm giá trị j tương ứng khóa "dài" L_{i_j}. Đó là khóa nằm trong bộ khóa nhái, mà TBTDL_Tổ đang thí nghiệm

Hàm Tim_j dùng phương pháp mã hóa khóa phiên giả KP (K Pseudo) để tìm ra chỉ số j, qua sự chênh lệch xác suất giải mã tM của TBTDL_Tổ , giữa hai lần mã hóa kề nhau. Khóa phiên giả KP có cùng độ dài với khóa phiên K.

Đặt p_j (j=0,…, m) là xác suất giải mã bản tin tM của TBTDL_Tổ khi ổ CCDL mã hóa j lần với khóa phiên giả KP và (m-j) lần với khóa đúng K, khi đó bản mã như sau [3]:

$$< [i_1,...,i_m, E(KP, L_{i_1}),..., E(KP, L_{i_{j-1}}), E(KP, L_{i_j}),$$

$$E(K, L_{i_{j+1}}),..., E(K, L_{i_m})], F_K(tM) >$$

ổ hư vậy, nếu TBTDL_Tổ giải mã đúng bản tin tM thì $p_0 = 1$, $p_m = 0$ (***2.1***).

Định lý:

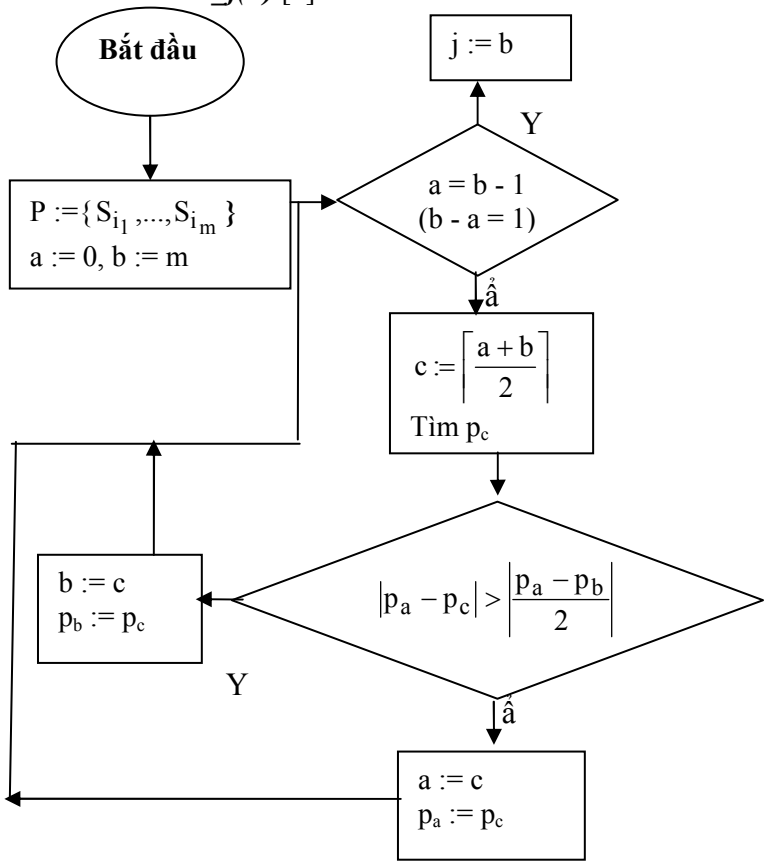
ổ ếu $|p_{j-1} - p_j| > \left|\frac{p_a - p_b}{2}\right|$ thì TBTDL_Tổ dùng khóa "dài" L_{i_j} của tập S_{i_j}.

Hàm Tim_j(P) tuân theo theo định lý này.

Với kết quả định lý trên, hàm Tim_j(P) dùng phương pháp tìm kiếm nhị phân để tìm cặp giá trị p_{j-1}, p_j giữa dãy giá trị $p_0,...,p_m$ thỏa mãn $|p_{j-1} - p_j| > \left|\frac{p_a - p_b}{2}\right|$.

Bắt đầu với khoảng [0, m], thu hẹp dần được khoảng [a, b]. Khi khoảng [a, b] chỉ còn chỉ số của một tập con, tức là $a = b - 1$, thì S_{i_b} chính là tập cần tìm, b chính là chỉ số j cần tìm trong hàm Tim_j(P).

Lưu đồ hàm Tim_j(P) [4]



2.3 LƯU VẾT VỚI NHIỀU BỘ KHÓA NHÁI

Phần 2.2 trình bày giải thuật lưu vết với một bộ khóa nhái. Trên thực tế, ả CCDL có thể thu mua được cùng lúc nhiều bộ khóa nhái. Để xử lý cùng lúc nhiều bộ khóa nhái, PM sẽ thực hiện lưu vết song song các TBTDL thí nghiệm với đầu vào P, R giống nhau [3].

Đầu tiên, các PM sẽ lấy P và R từ trong CSDL của ả CCDL. Khi các PM thực hiện lưu vết, nếu một PM xác định được P và R mới, PM này sẽ lưu P và R này vào trong CSDL và tạm ngừng tất cả các PM đang chạy.

Sau đó các PM này lại tiếp tục thực hiện lưu vết với P, R lấy từ CSDL chung của ả CCDL.

Quá trình cứ tiếp tục như vậy và kết thúc khi tất cả các TBTDL thí nghiệm đều không thể giải mã đúng bản tin thử nghiệm tM.

Chương 3: GIẢI PHÁP THU HỒI TBTDL BẤT HỢP PHÁP

3.1 MỘT SỐ KHÁI NIỆM

ả hắc lại bài toán trong mục 1.4:

Bài toán:

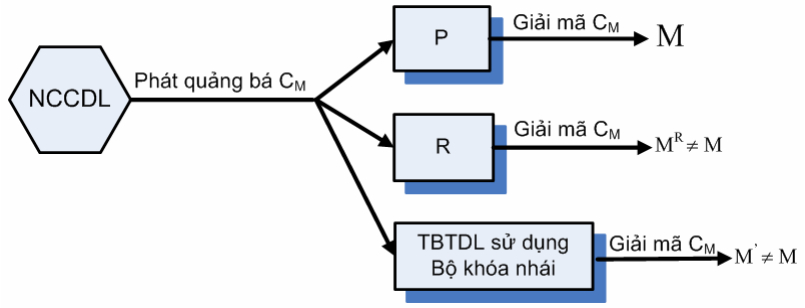
ả CCDL truyền thông điệp M qua các kênh quảng bá tới tập ả gồm các TBTDL (|ả |=n). Mỗi TBTDL có một tập các khóa. ả CCDL biết tập các TBTDL làm rò rỉ khóa R và tập P các TBTDL hợp pháp. P, R thỏa mãn: $P \cup R = ả$, $P \cap R = \phi$.

Yêu cầu:

Mọi TBTDL thuộc P phải giải mã được chính xác thông điệp M.

Mọi TBTDL thuộc R giải mã, chỉ thu được $M^R \neq M$.

Mọi TBTDL sử dụng bộ khóa nhái chỉ thu được $M' \neq M$.



Hình 3. 1: Mô hình thu hồi TBTDL bất hợp pháp

Mục đích chính của bài toán thu hồi là trừng trị những TBTDL bất hợp pháp. ả ghĩa là khi ả CCDL phát thông điệp M, những TBTDL bất hợp pháp sẽ chỉ giải mã được $M' \neq M$.

Để thực hiện được mục đích bài toán đặt ra, cần phải phân hoạch tập P thành các tập con $S_{i_1}, S_{i_2}, ..., S_{i_m}$, có các khóa "dài" tương ứng $L_{i_1}, L_{i_2}, ..., L_{i_m}$, trong đó các TBTDL $u \in S_{i_j}$ dùng chung khóa "dài" L_{i_j} , j=1, 2, …, m [3].

Khóa phiên K được mã hóa một lần với từng khóa $L_{i_1}, L_{i_2}, ..., L_{i_m}$, do đó chỉ những TBTDL nào thuộc một trong

các tập $S_{i_1}, S_{i_2}, ..., S_{i_m}$ mới có thể giải mã được K, và do đó giải mã được thông điệp M.

ả hư vậy, mọi TBTDL thuộc R và mọi TBTDL sử dụng bộ khóa nhái sẽ không giải mã chính xác được thông điệp M. ả ói cách khác, TBTDL bất hợp pháp sẽ không giải mã chính xác được thông điệp M.

Trong chương này, trình bày hai giải thuật phân hoạch tập P thành các tập con $S_{i_1}, S_{i_2}, ..., S_{i_m}$, là giải thuật **"Cây nhị phân con đầy đủ" (Complete Subtree - CS)** và giải thuật **"Hiệu hai tập con" (Subset Difference - SD)**.

Hai giải thuật CS và SD sử dụng "khung phủ tập con" (SCF - xem 1.2) để truyền thông điệp M tới TBTDL [3].

Cả hai giải thuật trên đều dựa trên cây nhị phân đầy đủ, với giả thiết số lượng TBTDL là $n = 2^k$, các TBTDL được biểu diễn là các lá của cây nhị phân, ký hiệu $u_1, u_2, .., u_n$. Các nút thuộc cây, kể cả lá được gán tên gọi $v_1, v_2, ..., v_{2n-1}$.

Giải thuật CS phân hoạch tập P thành các cây nhị phân con đầy đủ. Dùng nhãn tại gốc các cây con làm khóa "dài" để mã hóa khóa phiên K.

Giải thuật SD phân hoạch tập P thành tập $S_{i,j} = S_i - S_j$. Dùng bộ sinh số ngẫu nhiên với nguồn sinh là nhãn tại gốc cây con để sinh khóa "dài". Khóa "dài" này được dùng để mã hóa khóa phiên K.

Ký hiệu:

- S_i: tập các lá của cây nhị phân con gốc v_i.
- S_j: tập các lá của cây nhị phân con gốc v_j.
- $S_{i,j} = S_i - S_j$: chứa các lá thuộc S_i, nhưng không thuộc S_j; v_j là con của v_i.

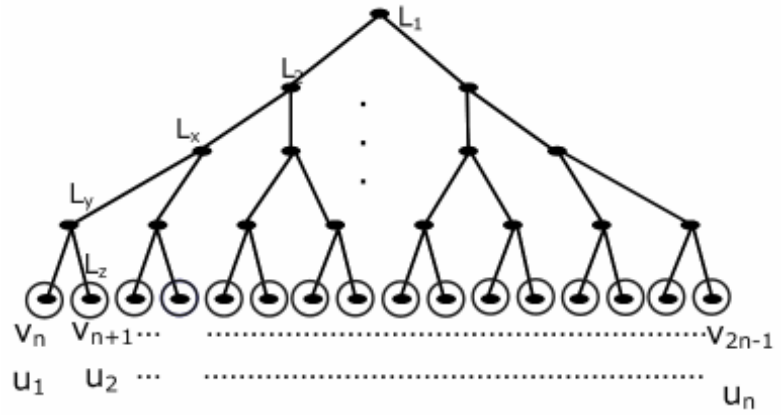
3.2 GIẢI THUẬT CÂY NHỊ PHÂN CON ĐẦY ĐỦ (Complete Subtree)

3.2.1 Ví dụ về giải thuật CS

3.2.2 Giải thuật CS

Giải thuật CS thực chất là giải thuật phân hoạch tập P các TBTDL hợp pháp thành tập các cây nhị phân con đầy đủ, và dùng các khóa "dài" tại gốc các cây con này, để mã hóa khóa phiên K [3].

Giả thiết có tập ả gồm n TBTDL, $n=2^k$, đã biết tập R các TBTDL làm rò rỉ khóa. ả CCDL biểu diễn các TBTDL thành cây nhị phân đầy đủ T, các TBTDL tương ứng với các lá của cây nhị phân, ký hiệu $u_1, u_2, .., u_n$. Các nút thuộc cây, kể cả lá được gán tên gọi $v_1, v_2, ..., v_{2n-1}$.



Hình 3.5: Cây nhị phân T biểu diễn n TBTDL

Phương pháp phân hoạch CS:

- ả CCDL xác định cây Steiner ST(R) bằng cách nối các lá thuộc R và gốc.

- ả CCDL tìm các nút còn lại trong cây nhị phân đầy đủ T, liền kề với các nút thuộc cây ST(R): được các nút $v_{i_1}, v_{i_2}, ..., v_{i_m}$.
- Ký hiệu S_{i_j} là tập các lá của cây con gốc v_{i_j}.
- $S_{i_1}, S_{i_2}, ..., S_{i_m}$ chính là các tập con của tập phủ P cần tìm: $P=\{S_{i_1}, S_{i_2}, ..., S_{i_m}\}$.
- Các nhãn $L_{i_1}, L_{i_2}, ..., L_{i_m}$ tương ứng với các nút $v_{i_1}, v_{i_2}, ..., v_{i_m}$ được dùng làm khóa "dài" cho các tập $S_{i_1}, S_{i_2}, ..., S_{i_m}$.

3.2.3 Hiệu năng của giải thuật CS

Trong giải thuật CS [3]:

- Độ dài phần đầu bản mã nhiều nhất là r.log(n) khóa (vì liên quan tới nhiều nhất r.log(n) khóa "dài").
- Số lượng các khóa "dài" của mỗi TBTDL là log(n).
- Thực hiện O(m.log(w)) phép toán + 1 phép giải mã thông điệp M.

3.3 GIẢI THUẬT HIỆU HAI TẬP CON (Subset Difference)

ả hược điểm của giải thuật CS là phân hoạch P thành một số lượng lớn các tập con. Giải thuật "Hiệu hai tập con" (Subset Difference - SD) sẽ khắc phục nhược điểm này bằng cách giảm số lượng các tập con [3].

Giải thuật SD cũng dựa trên cây, gần giống giải thuật CS, cũng sử dụng cây nhị phân để xác định các tập con của các TBTDL dùng chung các khóa "dài".

Trong giải thuật SD, ả CCDL phân hoạch tập P các TBTDL hợp pháp thành các tập $S_{i,j}$ sao cho: $S_{i,j} = S_i - S_j$. Trong đó:

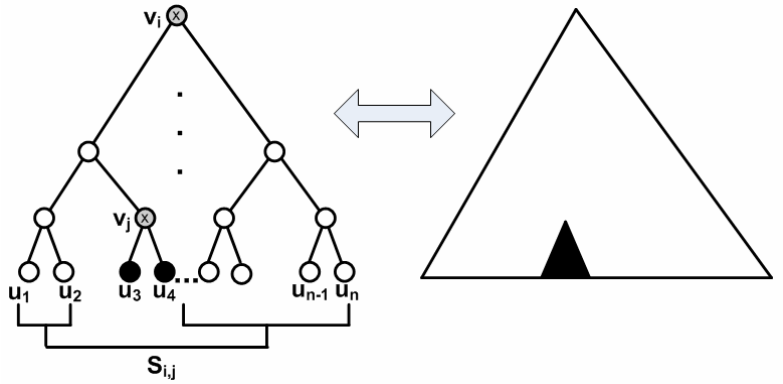
- S_i là tập các lá (TBTDL) của cây nhị phân gốc v_i.
- S_j là tập các lá (TBTDL bất hợp pháp) của cây nhị phân gốc v_j.

- v_i là cha của nút v_j.

Khóa "dài" của $S_{i,j}$ là $L_{i,j}$.

Các lá trong cây được gán nhãn $v_n,...,v_{2n-1}$, mỗi TBTDL tương ứng với một lá. Tập con $S_{i,j}$ các TBTDL hợp pháp là hiệu số: các lá của cây gốc v_i, trừ đi các lá của cây gốc v_j.

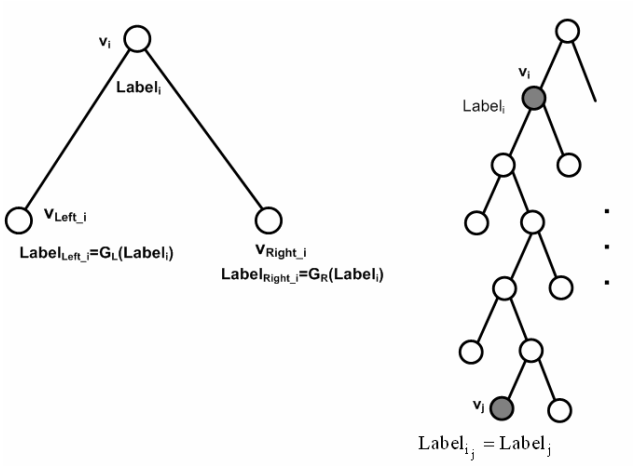
Vì v_i là cha v_j, nên $S_{i,j}$ chỉ có các lá của S_i, nhưng không có các lá của S_j.



Hình 3.6: $S_{i,j}$ chứa các lá thuộc cây nhị phân gốc v_i, nhưng không thuộc cây gốc v_j

3.3.1 Bộ sinh số ngẫu nhiên G

Gọi $G:\{0,1\}^k \rightarrow \{0,1\}^{3k}$ là bộ sinh số ngẫu nhiên sinh 3 giá trị G_L, G_R, G_M. Xét mầm sinh S=$Label_i$ (seed) là nhãn tại nút v_i, các giá trị có thể sinh từ G là: $G_L(Label_i)$, $G_R(Label_i)$, $G_M(Label_i)$, trong đó: $G_L(Label_i)$ là nhãn tại nút là con trái của v_i, $G_R(Label_i)$ là nhãn tại nút là con phải của v_i. $G_M(Label_i)$ là khóa tại nút v_i [3].



Hình 3.7: Bộ sinh số ngẫu nhiên G với mầm sinh $Label_i$

Ký hiệu $Label_{i,j}$ là nhãn tại nút v_j : $Label_{i,j} = Label_j$, v_j là con hoặc cháu của v_i . $Label_i$ là nhãn tại nút v_i .

$Label_j$ là nhãn tại nút v_j, được tính theo bộ sinh số ngẫu nhiên G, với mầm sinh $Label_i$, dựa trên đường đi từ v_i xuống v_j .

Ký hiệu $L_{i,j}$ là khóa "dài" của tập $S_{i,j}$, $L_{i,j} = G_M(Label_{i,j})$.

3.3.2 Ví dụ về giải thuật SD

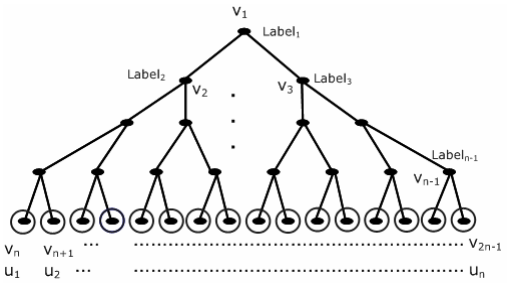
3.3.3 Giải thuật SD

Giải thuật SD thực chất là giải thuật phân hoạch tập P các TBTDL hợp pháp thành các tập $S_{i,j}$, $S_{i,j} = S_i - S_j$. Trong đó:

- S_i là tập các lá (TBTDL) của cây nhị phân gốc v_i .
- S_j là tập các lá (TBTDL bất hợp pháp) của cây nhị phân gốc v_j.
- v_i là cha của nút v_j.

Giải thuật SD dùng các khóa "dài" là hàm của nhãn tại gốc các nút v_i để mã hóa khóa phiên K [3].

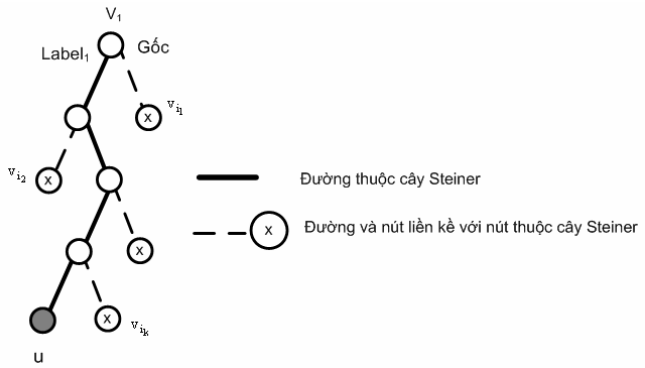
Giả thiết có tập ầ gồm n TBTDL, $n=2^k$, đã biết tập R các TBTDL làm rò rỉ khóa. ầ CCDL biểu diễn các TBTDL thành cây nhị phân đầy đủ, các TBTDL tương ứng với các lá của cây nhị phân, ký hiệu $u_1, u_2, .., u_n$. Các nút thuộc cây, kể cả lá được gán tên gọi $v_1, v_2, ..., v_{2n-1}$.



Hình 3.15: Cây T biểu diễn toàn bộ n TBTDL

Giải thuật thiết lập khóa:

- Xác định cây Steiner nối u_i với gốc, ký hiệu ST(u_i). Đặt V(ST(u_i)) là các nút thuộc ST(u_i), trừ nút u_i.
- Tìm các nút (kể cả lá) liền kề với các nút thuộc V(ST(u_i)), ký hiệu KeV(ST(u_i)). Các nút này thuộc cây T, nhưng không thuộc ST(u_i).



Hình 3.16: Minh họa cây ST(u) và các nút KeV(ST(u))

- **Sinh khóa cho u_i:**

  $L_{u_i} = \{Label_{i,j}\}$, trong đó:

  $v_i \in V(ST(u_i))$, $v_j \in KeV(ST(u_i))$.

  v_j là con, hoặc cháu của v_i.

Phân hoạch P thành các tập con:

- Xác định cây Steiner ST(R): nối gốc với R.
- Khởi tạo S={}, T= ST(R).

+ **Bước 1**: Tìm hai lá (v_i, v_j) trong T, tìm v là ***cha chung thấp nhất*** của (v_i, v_j). Đặt v_h là cha riêng của v_i, v_k là cha riêng của v_j.

+ **Bước 2**:

  ẩ ếu $v \neq v_h \neq v_i$ thì $S = S \cup S_{h,i}$; ẩ ếu $v \equiv v_h \neq v_i$ thì S không thay đổi;

  ẩ ếu $v \neq v_k \neq v_j$ thì $S = S \cup S_{k,j}$; ẩ ếu $v \equiv v_k \neq v_j$ thì S không thay đổi;

+ **Bước 3**: T = T - {các con của v}, đánh dấu v là lá.
  Tức là giữ lại cha chung thấp nhất v.

- Lặp lại các bước từ 1 đến 3 cho đến khi T chỉ còn 1 nút.
  ẩ ếu còn $v_i \neq$ gốc thì $S=S \cup S_{gốc,\, i}$.
  Kết thúc được tập $S = \{S_{i_1,j_1}, S_{i_2,j_2}, ..., S_{i_m,j_m}\}$.

Các tập $S_{i_1,j_1}, S_{i_2,j_2}, ..., S_{i_m,j_m}$ có các khóa "dài" $L_{i_1,j_1}, L_{i_2,j_2}, ..., L_{i_m,j_m}$ được sinh dựa trên bộ sinh số ngẫu nhiên G:

$$L_{i_1,j_1} = G_M(Label_{i_1,j_1})$$
$$L_{i_2,j_2} = G_M(Label_{i_2,j_2})$$
$$....$$
$$L_{i_m,j_m} = G_M(Label_{i_m,j_m})$$

3.3.4 Hiệu năng của giải thuật SD

Trong giải thuật SD [3]:

- Độ dài phần đầu bản mã nhiều nhất là 2r-1 khóa.
- Số lượng các khóa "dài" của mỗi TBTDL nhiều nhất là $\frac{1}{2}\log^2 n + \frac{1}{2}\log n + 1$.
- Thực hiện O(m.log(w)) phép toán + 1 phép giải mã thông điệp M.

Chương 4: ĐỘ AN TOÀN CỦA GIẢI THUẬT KHUNG PHỦ TẬP CON

4.1 CÀI ĐẶT GIẢI THUẬT E MÃ HÓA KHÓA PHIÊN K

4.1.1 Phương pháp "Cắt phần đầu bản mã" (Prefix Truncation)

4.1.2 Phương pháp mã hóa khóa công khai cho E

4.2 CÀI ĐẶT GIẢI THUẬT F MÃ HÓA BẢN TIN M

4.3 ĐỘ AN TOÀN CỦA GIẢI THUẬT SCF

4.3.1 Độ an toàn của giải thuật E, F và giải thuật thiết lập khóa.

Trong phần 4.3, thuật ngữ ***TBTDL bất hợp pháp*** nghĩa là ***TBTDL này*** sử dụng ***bộ khóa nhái.***

1) Độ an toàn giải thuật E

TBTDL bất hợp pháp B nghiên cứu hệ mã hóa E, bản mã C.

Có hai trường hợp khi TBTDL B nhận được bản mã của khóa phiên:

(i) Hoặc $E_L(K)$.

(ii) Hoặc $E_L(RK)$, trong đó RK (Random K) là chuỗi ngẫu nhiên có độ dài |K|.

TBTDL B phân biệt được hai trường hợp trên với xác suất không đáng kể.

Hiệu số của chúng giới hạn bởi ε_1:

$$|\Pr[\text{B giải mã } E_L(K)] - \Pr[\text{B giải mã } E_L(RK)]| \le \varepsilon_1 \qquad (4.1)$$

Trong đó: Pr ký hiệu xác suất giải mã E.

$\Pr[\text{B giải mã } E_L(K)]$ là xác suất giải mã $E_L(K)$ đúng của B.

$\Pr[\text{B giải mã } E_L(RK)]$ là xác suất giải mã $E_L(RK)$ đúng của B.

ả hư vậy giải thuật E có độ an toàn $\le \varepsilon_1$ [3].

2) Độ an toàn giải thuật F

Có hai trường hợp khi TBTDL B bất hợp pháp nhận được bản mã của bản tin M:

(i) Hoặc $F_K(M)$.

(ii) Hoặc $F_K(RM)$, trong đó RM là chuỗi ngẫu nhiên có độ dài |M|.

TBTDL B phân biệt được 2 trường hợp với xác suất không đáng kể.

Hiệu số của chúng giới hạn bởi ε_2:

$$|\Pr[\text{B giải mã } F_K(M)] - \Pr[\text{B giải mã } F_K(RM)]| \le \varepsilon_2 \qquad (4.2)$$

Trong đó: Pr ký hiệu xác suất giải mã F.

$\Pr[\text{B giải mã } F_K(M)]$ là xác suất giải mã $F_K(M)$ đúng của B.

$\Pr[\text{B giải mã } F_K(RM)]$ là xác suất giải mã $F_K(RM)$ đúng của B.

ả hư vậy giải thuật F có độ an toàn $\le \varepsilon_2$ [3].

3 Độ an toàn của giải thuật thiết lập khóa

Trong SCF còn có giải thuật thiết lập khóa "dài" L_i cho mỗi TBTDL $u \in S_i$.

Định nghĩa 4.4

Gọi A là giải thuật sử dụng SCF xác định các tập con $S_1,...,S_w$. Xét TBTDL bất hợp pháp B:

- Lựa chọn i, $1 \le i \le w$.
- ả hận tập khóa bí mật L_u cho $\forall u \in$ ả $- S_i$

Ta nói rằng A thỏa mãn thuộc tính "**khóa không thể phân biệt được**", nếu hiệu số xác suất B giải mã L_i và xác suất B giải mã RL_i (RL_i là chuỗi ngẫu nhiên cùng độ dài L_i) là không đáng kể và biểu diễn bởi ε_3:

$$|\Pr[\text{B giải mã } L_i] - \Pr[\text{B giải mã } RL_i]| \le \varepsilon_3 \qquad (4.3)$$

Trong đó: Pr ký hiệu xác suất giải mã khóa "dài"..

$\Pr[\text{B giải mã } L_i]$ là xác suất giải mã L_i với kết quả đúng của B.

$\Pr[\text{B giải mã } RL_i]$ là xác suất giải mã RL_i với kết quả đúng của B.

ả hư vậy giải thuật thiết lập khóa có độ an toàn $\leq \varepsilon_3$ [3].

Bổ đề tiếp theo [3] là hệ quả của thuộc tính "**khóa không thể phân biệt được**", và sẽ được sử dụng để chứng minh định lý 4.1.

Bổ đề 4.1

Với bất kỳ $1 \leq i \leq w$, đặt $S_{i_1}, S_{i_2}, ..., S_{i_t}$ là các tập con chứa trong S_i.

Đặt $L_{i_1}, L_{i_2}, ..., L_{i_t}$ là các khóa tương ứng với các tập con đó.

TBTDL bất hợp pháp B chọn i, $1 \leq i \leq w$, và nhận L_u cho $\forall u \in$ ả $- S_i$.

ả ếu B chú ý phân biệt các khóa $L_{i_1}, L_{i_2}, ..., L_{i_t}$ với các khóa ngẫu nhiên $RL_{i_1}, RL_{i_2}, ..., RL_{i_t}$ thì:

$$|\Pr[B \text{ giải mã } L_{i_1}, ..., L_{i_t}] - \Pr[B \text{ giải mã } RL_1, .., RL_{i_t}]| \leq t.\ \varepsilon_3 \quad (4.4)$$

Trong đó: Pr ký hiệu của xác suất giải mã khóa "dài".

$\Pr[B$ giải mã $L_{i_1}, L_{i_2}, ..., L_{i_t}]$ là xác suất giải mã khóa phiên K, được mã hóa lần lượt t lần với từng $L_{i_1}, L_{i_2}, ..., L_{i_t}$.

$\Pr[B$ giải mã $RL_{i_1}, RL_{i_2}, ..., RL_{i_t}]$ là xác suất giải mã khóa phiên K, được mã hóa lần lượt t lần với từng $RL_{i_1}, RL_{i_2}, ..., RL_{i_r}$.

4.3.2 Khái niệm độ an toàn của giải thuật SCF

Để định nghĩa độ an toàn của giải thuật SCF, ta xem xét TBTDL bất hợp pháp trong ngữ cảnh sau: TBTDL bất hợp pháp có thể có toàn bộ khóa hoặc một phần bộ khóa của một số TBTDL, và nó có thể có một vài tác động trên sự lựa chọn các thông điệp được mã hóa (lựa chọn bản gốc). Kẻ gian mạo danh ả CCDL có thể tạo các thông điệp giả và quan sát sự phản ứng của các TBTDL hợp pháp [3].

Độ an toàn của giải thuật SCF là một hàm về độ an toàn của E và F. Đo độ an toàn bằng xác suất.

Định nghĩa 4.5

Xét TBTDL bất hợp pháp B nhận được:

- Lựa chọn thích hợp tập R các TBTDL và thu được L_u với mọi u ∈ R.

B mạo danh ả CCDL lựa chọn các thông điệp M_1, M_2,… và tập các TBTDL bị thu hồi R_1, R_2,…(các tập này không cần thiết phải tương ứng với các tập TBTDL cần thu hồi thực sự) và quan sát sự mã hóa M_i khi tập các TBTDL bị thu hồi là R_i.

B có thể tạo ra bản mã, và quan sát TBTDL u hợp pháp giải mã. B có thể thám khóa của TBTDL u và thu được L_u. Bước tiếp theo lặp lại |R| lần với các u ∈ R.

- Chọn thông điệp gốc M ngẫu nhiên và tập R các TBTDL cần thu hồi.

Thì B sẽ nhận được bản rõ M' với tập các TBTDL cần thu hồi R.

ả ó phải phỏng đoán phải chăng M'= M (a), hoặc M'= RM (b), trong đó RM là thông điệp ngẫu nhiên có cùng độ dài với M.

Giải thuật thu hồi là **an toàn,** nếu với bất kỳ TBTDL bất hợp pháp B ở trên, hiệu số giữa xác suất giải mã đúng M của B và xác suất giải mã đúng RM của B là không đáng kể.

Định lý 4.1 (về độ an toàn của giải thuật SCF) [3]

Cho A là giải thuật dùng SCF, trong đó giải thuật thiết lập khóa thỏa mãn thuộc tính "**khóa không thể phân biệt được**". E thỏa mãn biểu thức 4.1, F thỏa mãn biểu thức 4.2.

Thì A là **an toàn** theo định nghĩa 4.5 với **hệ số an toàn** δ, $\delta \le \varepsilon_2 + 2mw(\varepsilon_1 + 4w\varepsilon_3)$, trong đó w là số lượng tập con S_{i_j} trong giải thuật và m là kích thước lớn nhất của phủ P,

δ=|Pr[B giải mã M] - Pr[B giải mã RM]|.

Với: Pr ký hiệu xác suất.

Pr[B giải mã M] là xác suất giải mã thông điệp M của B.

Pr[B giải mã RM] là xác suất giải mã thông điệp ngẫu nhiên RM của B.

Chương 5: MỘT SỐ ỨNG DỤNG

5.1 TRUYỀN HÌNH INTERNET

5.2 TRUYỀN HÌNH DI ĐỘNG

KẾT LUẬN

ả hiệm vụ của luận văn là trình bày giải pháp lưu vết và thu hồi TBTDL bất hợp pháp. Kết quả chính của luận văn là nghiên cứu tài liệu để hệ thống lại các vấn đề sau:

1) Bài toán lưu vết: luận văn đã trình bày giải pháp lưu vết tập con để xác định được định danh TBTDL làm rò rỉ khóa qua phương pháp tìm kiếm nhị phân.
2) Bài toán thu hồi: luận văn đã trình bày hai giải pháp để phân hoạch tập các TBTDL thành các tập con dùng chung khóa "dài". Đó là giải pháp "Cây nhị phân con đầy đủ" (Complete Subtree) và "Hiệu hai tập con" (Subset Difference).

Trên thực tế có nhiều giải pháp để thu hồi và lưu vết kẻ phạm tội, tuy nhiên luận văn chọn hướng nghiên cứu các giải pháp này, do chúng đã được đưa vào chuẩn của những nước phát triển mạnh trên thế giới là châu Âu (tháng 2/2002) và Mỹ (tháng 7/2007).

3) Độ an toàn của giải thuật Subset Cover Framework.
4) Một số ứng dụng.

Lĩnh vực bảo mật trong việc truyền dữ liệu quảng bá từ ả CCDL đến các TBTDL còn rất nhiều hướng đi và các phần cần nghiên cứu. ả hững kiến thức luận văn đã trình bày chỉ là rất nhỏ so với thực tế cần. Tác giả luận văn mong muốn nhận được sự quan tâm và góp ý của các thầy, cô giáo, bạn bè, những chuyên gia trong lĩnh vực này để luận văn hoàn thiện hơn, và có thể đóng góp vào công tác bảo vệ bản quyền các sản phẩm số hóa tại Việt ả am trong tương lai gần.

TÀI LIỆU THAM KHẢO

TIẾNG VIỆT

1. Gs Phan Đình Diệu (2002), "*Lý thuyết mật mã & An toàn thông tin*", Đại học Quốc gia Hà ả ội, tr 73 – 75.

TIẾNG ANH

2. Benny Chor, Amos Fiat, Moni ả aor, Benny Pinkas (02/05/2000), "*Tracing Traitors*", IEEE Transactions on Informatin Theory, Volume 46.
3. Dalit ả aor, Moni ả aor, Jeff Lostpiech (24/02/2001), "*Revocation and Tracing Schemes for Stateless Receivers*", Advances in Cryptology – Crypto '01 (Berlin), Lecture ả otes in Computer Science Volume 2139.
4. Lotspiech et al (07/03/2006), "*Method for tracing traitor receivers in a broadcast encryption system*", United States Patent, ả o 7070125 B2, pp 3 – 10.
5. Ronald L. Rivest (1997), "*All-or-Nothing Encryption and The Package Transform*". Proc. 4th Fast Software Encryption International Workshop, Lecture ả otes in Computer Science, Volume 1174.
6. Thomas Martin (05/042005), "*A set theoretic approach to broadcast encryption*". Royal Holloway University of London.
7. Yevgeniy Dodis, ả elly Fazio (01/082002), "*Public Key Broadcast Encryption for Stateless Receivers*", Digital Right Management, - DRM '02, Lả CS 2696.
8. Zbigniew J. Czech, George Havas, Bohdan S. Majewski (1992), "*An optimal algorithm for generating minimal perfect hash function*", Information Processing Letter.