Bảo mật dữ liệu tầng vật lý trong mạng truyền tin không dây: Những ý tưởng đầu tiên và hướng nghiên cứu hiện nay

Đặng Vũ Sơn và Nguyễn Như Tuấn, Ban Cơ yếu Chính phủ

Bên cạnh việc áp dụng các kỹ thuật mã hóa truyền thống tại các tầng trên (trong các mô hình truyền tin phân tầng) để bảo mật dữ liệu, ý tưởng về bảo mật tại tầng vật lý cho mạng truyền tin không dây đã được đề cập từ những năm 1970 và cho đến nay, đặc biệt trong một thập kỷ gần đây, thì ý tưởng này đang được cộng đồng các nhà nghiên cứu khoa học trên toàn thế giới quan tâm.

Bài báo này sẽ giới thiệu các kết quả đầu tiên về bảo mật tầng vật lý cho mạng truyền tin không dây dựa trên 02 công trình, một của Wyner ([1]) về kênh nghe lén (The wire-tap channel) được công bố vào năm 1975 và một của Imre Csiszár và János Körner ([2]) về các kênh truyền quảng bá với các thông báo bí mật (the broadcast channels with confidential messages) được công bố vào năm 1978. Bài báo cũng giới thiệu ngắn gọn một số kết quả và hướng nghiên cứu này trong thời gian gần đây.

1 Giới thiệu

Hiện nay, hầu hết các phương pháp đảm bảo bí mật trong hệ thống truyền tin là dựa vào kỹ thuật mật mã để mã hóa nội dung thông tin cần bảo mật từ nơi gửi đến nơi nhận. Chúng ta cùng xem xét một mô hình truyền tin cơ bản như Hình 1.

Người gửi là Alice muốn gửi một thông báo cho người nhận là Bob. Còn Eve - người nghe lén, không thể biết được nội dung thông báo. Để đảm bảo yêu cầu trên, Alice sử dụng một

Hình 1: Mô hình hệ thống truyền tin với Alice là người gửi, Bob là người nhận hợp pháp và Eve là kẻ nghe trộm (Eavesdropper).

hoặc nhiều thuật toán mã hóa bảo mật kết hợp với khóa mã để mã hóa bản thông báo. Bob biết về thuật toán mã hóa được sử dụng nên đã dùng khóa bí mật hợp lệ do anh ta có để giải mã bản thông báo. Còn Eve, có thể biết về thuật toán mã hóa được sử dụng, nhưng không biết về khóa mã được sử dụng, nên sẽ rất khó có thể giải mã được thông báo do Alice gửi cho Bob.

Một xu hướng khác mới nổi trong bảo mật mạng không dây trong thời gian gần đây là bảo mật dữ liệu truyền tin từ tầng vật lý (Physical Layer Security - PLS) mà không dùng mật mã. Hướng nghiên cứu này được khởi xướng từ năm 1975 bởi Tiến sĩ Aaron D. Wyner ([1]), một nhà khoa học người Mỹ. Trong công trình này, Wyner đã chứng minh rằng có thể truyền tin bảo mật với tốc độ C_s ($C_s > 0$) trên kênh truyền có sự xuất hiện của người nghe lén (Eavesdropper). Một giả thiết quan trọng trong các kết quả của Wyner là kênh truyền giữa người gửi (Alice) và người nghe lén (Eve), sau đây gọi tắt là kênh nghe lén (wire-tap channel), có độ suy hao lớn hơn kênh truyền từ gửi (Alice) đến người nhận hợp pháp (Bob), sau đây gọi là kênh chính (main channel). Phần 2 của bài báo sẽ giới thiệu các kết quả chính của công trình này.

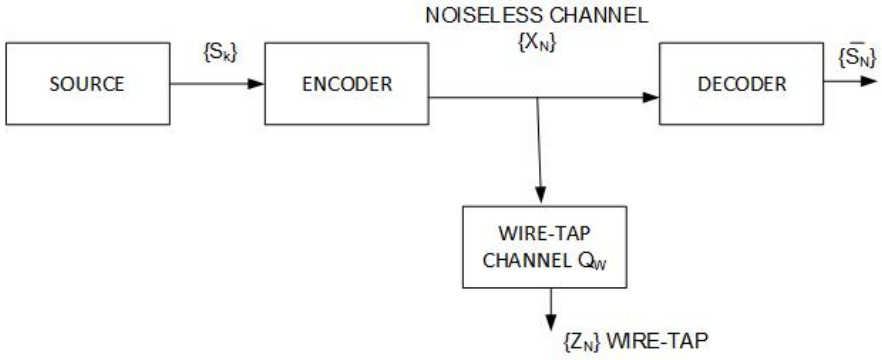
Một phát triển mở rộng hơn cho các kết quả của Aaron D. Wyner được công bố bởi hai nhà khoa học người Hungari là Imre Csiszár và János Körner vào năm 1978 ([2]) là có thể truyền thông báo bí mật (confidential message) tại tốc độ C_s ($C_s > 0$) với độ bảo mật hoàn hảo (perfect secrecy) cùng với thông báo chung (common message) không cần giữ bí mật cho tất cả mọi người trong hệ thống. Tất cả các kết quả trên đều xuất phát từ bài toán tối ưu trong lý thuyết truyền tin (Information theory). Phần 3 sau đây sẽ giới thiệu các kết quả này. Phần 4 của bài báo sẽ giới thiệu tóm tắt một số nghiên cứu về lĩnh vực này trong thời gian gần đây và kết luận.

2 Kênh nghe lén

Bài báo "The Wire-tap channel" được coi là một công trình lớn nhất của A. D. Wyner trong gần 20 năm (1974 - 1993), khi ông ở vị trí trưởng bộ phận Nghiên cứu phân tích truyền thông (communications analysis research) của tập đoàn truyền thông Bell tại Murray Hill, New Jersey, Mỹ. Tuy nhiên, khoảng hơn 30 năm sau khi bài báo được công bố thì các kết quả này mới thực sự được các nhà khoa học trên khắp thế giới tập trung nghiên cứu và phát triển.

Bài báo quan tâm đến hệ thống truyền tín hiệu số trên kênh rời rạc, không nhớ (Descrete, Memoryless Channel - DMC) có nhiễu và có sự tham gia của người nghe lén (wire-tapper) tại một kênh DMC có nhiễu khác, như Hình 1. Trong đó, Alice và Bob có sử dụng các kỹ thuật mã hóa và giải mã, tuy nhiên phương thức mã hóa được giả thiết là cũng được biết bởi Eve.

Kênh truyền trong hệ thống cũng được giả thiết là hoàn hảo, không có lỗi (perfect transmission, error-free). Bài báo đã chỉ ra được quan hệ trong cặp giá trị (R, d), với R là tốc độ truyền tin cực đại từ Alice tới Bob, d là độ mập mờ (equivocation) về nguồn tin của người nghe lén (Eve) đối với dữ liệu thu được. Đặc biệt, nếu d bằng với độ bất định (entropy) của nguồn tin H_s thì chúng ta kết luận rằng quá trình truyền tin là tuyệt đối an toàn. Kết luận của Wyner chứng tỏ rằng, tồn tại giá trị $C_s > 0$, theo đó quá trình truyền tin tin cậy có thể



Hình 2: Kênh nghe lén đặc biệt

đạt tới tốc độ C_s (secrecy capacity) là có thể chấp nhận như tuyệt đối an toàn.

2.1 Mô hình truyền tin và phát biểu bài toán

Ban đầu, bài toán được xem xét trên hệ thống truyền tin đặc biệt đơn giản như Hình 2. Bộ phát (the source) phát định hướng một chuỗi dữ liệu $S_1, S_2, \ldots$ là các bit độc lập, ngẫu nhiên S, $(Pr\{S=0\} = Pr\{S=1\} = \frac{1}{2})$. Bộ mã hóa nguồn kênh kiểm tra K bit nguồn đầu tiên $S^K = (S_1, \ldots, S_K)$ và mã hóa S^K sang vector nhị phân có độ dài N là $X^N = (X_1, \ldots, X_N)$. X^N được truyền lần lượt đến bộ giải mã thông qua kênh không nhiễu và được chuyển đổi thành dòng dữ liệu nhị phân $\bar{S^K} = (\bar{S_1}, \ldots, \bar{S_K})$ tại nơi nhận. Xác suất lỗi (error probability) trong trường hợp này được xác định như sau:

$$P_e = \frac{1}{K}\Sigma_{k=1}^{K} Pr\{S_k \neq \bar{S_k}\} \tag{1}$$

Toàn bộ quá trình xử lý được lặp lại cho đến khi truyền hết khối tin K . Tỷ lệ truyền tin là $\frac{K}{N}$ bit trên mỗi ký tự được truyền.

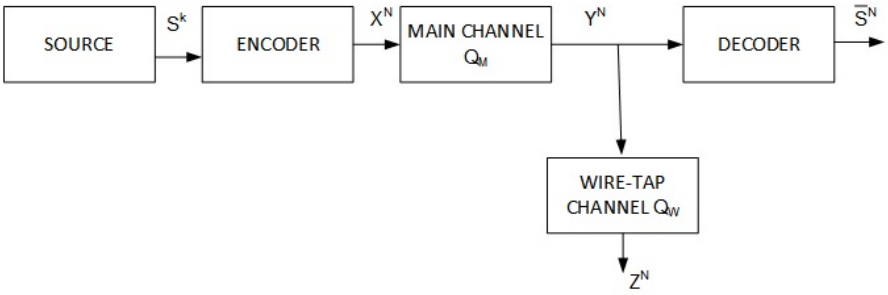
Người nghe lén thu được khối dữ liệu $Z^N = (Z_1, \cdots, Z_N)$ thông qua kênh nhị phân đối xứng (Binary symmetric channel - BSC) với xác suất chuyển đổi (crossover probability) p_0, $(0 < p_0 \leq \frac{1}{2})$, do vậy, với $x, z = 0, 1$, $(1 \leq n \leq N)$ thì:

$$Pr\{Z_n = z | X_n = x\} = (1 - p_0)\delta_{x,z} + p_0(1 - \delta_{x,z}).$$

Trong đó, $\delta_{x,z} = Q_W(z|x)$ là xác suất chuyển trên kênh nghe lén (wire-tap) rời rạc, không nhớ. Chúng ta có độ mập mờ về nguồn tin của người nghe lén, hay còn gọi là độ khó của việc xác định nguồn tin đã gửi tương ứng với dữ liệu đã nhận, được định nghĩa như sau :

$$\Delta \triangleq \frac{1}{K} H(S^K | Z^N) \tag{2}$$

Người thiết kế hệ thống sẽ mong muốn có được giá trị P_e tiến sát về 0, trong khi tỷ lệ K/N và giá trị Δ càng lớn càng tốt. Bài báo [1] đã chỉ ra rằng khi $N \to \infty$ thì sự mập mờ (equivocation) tại kênh nghe lén sẽ đạt entropy nguồn vô điều kiện, do vậy quá trình truyền tin là tuyệt đối an toàn. Nhưng khi $N \to \infty$ thì tốc độ truyền tin $\frac{K}{N} = \frac{1}{N} \to 0$. Vậy, câu hỏi



Hình 3: Kênh nghe lén trong trường hợp tổng quát

đặt ra là: liệu có thể truyền tin ở một tốc độ giới hạn lớn hơn 0 một lượng đáng kể mà vẫn đạt được mức độ an toàn gần như tuyệt đối ($\Delta \approx H(S_1)$).

2.2 Kết quả của A. D. Wyner

Câu hỏi đặt ra ở đoạn cuối của phần trên được trả lời trực tiếp qua mô hình truyền tin tổng quát như Hình 3.

Trong đó, nguồn tin gửi là rời rạc, không nhớ với entropy H_s. Kênh chính (main channel) và Kênh nghe lén (wire-tap channel) là kênh rời rạc, không nhớ có xác suất chuyển tương ứng là $Q_M(.|.)$ và $Q_W(.|.)$. Nguồn tin và xác suất chuyển Q_M và Q_W được cho trước và cố định. Bộ mã hóa nguồn hoạt động như là một kênh với dữ liệu đầu vào là vector có độ dài K (S^K) và đầu ra là vector có độ dài N (X^N). Vector X^N được đưa lần lượt vào kênh chính. Đầu ra của kênh chính và cũng là đầu vào của kênh nghe lén là vector Y^N. Đầu ra của kênh nghe lén là vector Z^N. Bộ giải mã nguồn tính toán ra vector $\bar{S^K}$ từ Y^N và xác suất lỗi P_e được cho bởi công thức (1). Độ mập mờ (equivocation) Δ như công thức (2) và tốc độ truyền tin là KHs/N bit nguồn trên một đơn vị đầu vào của kênh.

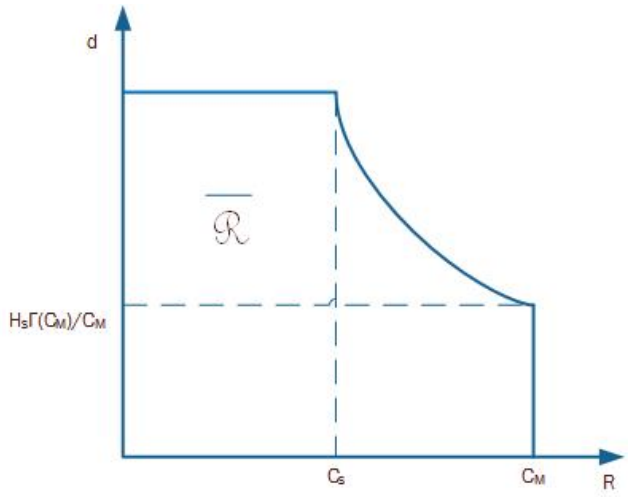
Wyner đã phát biểu rằng, cặp giá trị (R, d) có thể đạt được nếu tìm ra bộ mã hóa – giải mã nguồn với sự biến đổi nhỏ của P_e, tốc độ truyền KHs/N tương đương với R và độ mập mờ Δ tương đương với d (với N và K có thể là rất lớn). Kết quả chính của bài toán là đã tìm ra các đặc trưng của họ các cặp (R, d) đạt được như thể hiện trên Hình 4, với miền $\bar{\mathcal{R}}$ được xác định như sau

$$\bar{\mathcal{R}} \triangleq \{(R,d):\ 0 \le R \le C_M,\ 0 \le d \le H_s,\ R_d \le H_s\Gamma(R)\} \tag{3}$$

với

$$\Gamma(R) \triangleq \sup_{p_x \in \mathcal{P}(R)} I(X;Y|Z) = \sup_{p_x \in \mathcal{P}(R)} [I(X;Y) - I(X;Z)]$$

Trong đó, $I(A;B)$ là lượng thông tin tương hỗ (mutual information) giữa A và B; hàm *sup* (supremum) tương đương với hàm *max*; $p_x(x) = P_r\{X = x\},\ x \in \mathcal{X}$ và $\mathcal{P}(R)$ là tập các giá trị của p_x sao cho $I(X;Y) \ge R$; C_M là dung lượng kênh chính.



Hình 4: Miền giá trị của $\bar{\mathcal{R}}$

Theo như các đặc trưng được đưa ra trong (3) hoặc như mô tả trên Hình 4, gần như trong tất cả các trường hợp, luôn tồn tại một giá trị "secrecy capacity" $C_s > 0$. Theo đó cặp giá trị (R, d) tương đương với (C_s, H_s) là có thể đạt được [trong khi nếu $R > C_s$, thì (R, H_s) là không thể đạt được]. Do vậy, có thể truyền tin ở tốc độ C_s với độ an toàn gần như tuyệt đối.

3 Kênh truyền quảng bá với các thông báo bí mật

Tiếp theo nghiên cứu của Aaron D. Wyner, hai nhà khoa học người Hungari cùng làm việc tại Viện Toán học thuộc Học viện Khoa học Hungari (Mathematical Institute of the Hungarian Academy of Sciences) là Imre Csiszár và János Körner đã công bố một kết quả phát triển hơn vào năm 1978 ([2]).

Kết quả của Imre Csiszár và János Körner được phát biểu trên hệ thống truyền tin như Hình 1, trong trường hợp truyền tin quảng bá với bộ 3 tham số đặc trưng là (R_1, R_e, R_0). Theo đó, Alice truyền quảng bá thông báo chung (common message) cho cả Bob và Eve với cùng một tốc độ là R_0. Bên cạnh đó thì Alice lại truyền một thông báo riêng (private message) đến Bob với tốc độ là R_1. Cả Alive và Bob không muốn Eve biết nội dung thông báo riêng này, độ mập mờ (equivocation) tối thiểu của Eve đối với thông báo riêng là R_e. Mô hình của Imre Csiszár và János Körner khác với mô hình đề cập trong [1] ở hai điểm cơ bản. Thứ nhất là Imre Csiszár và János Körner không đưa ra giả thiết là kênh truyền từ Alice đến Eve là yếu hơn kênh truyền từ Alice tới Bob. Thứ hai là trong công trình của A. D. Wyner không đề cập đến trường hợp Alice truyền thông báo chung theo cách quảng bá đến cả Bob và Eve.

3.1 Phát biểu bài toán

Trong hệ thống kênh truyền quảng bá với các thông báo bí mật BCC (Broadcast Channels with Confidential messages), một bộ mã khối tất định f là một ánh xạ $f : \mathcal{S} \times \mathcal{T} \to \mathcal{X}^n$, với

$\mathcal{S}$ và $\mathcal{T}$ là hai tập tùy biến tương ứng biểu diễn cho các thông báo chung và thông báo riêng có thể. Định nghĩa về bộ mã khối f được định nghĩa trong [2] như sau:

Định nghĩa 1. *Một bộ mã khối f với độ dài khối là n trong hệ thống BCC được xác định bởi một ma trận các xác suất điều kiện (matrix of conditional probabilities) $f(x^n|s,t)$. Trong đó, $x^n \in \mathcal{X}^n,\ s \in \mathcal{S},\ t \in \mathcal{T},\ \Sigma_{x^n} f(x^n|s,t) = 1$ và $f(x^n|s,t)$ là xác suất để cặp thông báo (s,t) được mã hóa thành đầu vào kênh x^n.*

Hai bộ giải mã, tại Bob và Eve, tương ứng với một cặp ánh xạ $\varphi : \mathcal{Y}^n \to \mathcal{X} \times \mathcal{T}$ và $\psi : \mathcal{Z}^n \to \mathcal{T}$. Khi này, hoạt động truyền tin trong hệ thống BCC được thực hiện theo bộ mã hóa - giải mã (f, φ, ψ) sao cho giá trị lỗi là tối ưu nhất.

Định nghĩa 2. *Bộ mã hóa - giải mã (f, φ, ψ) được gọi là $(n, \epsilon) - transmission$ trên BCC khi và chỉ khi với $s \in \mathcal{S}$, và $t \in \mathcal{T}$, bộ giải mã φ cho kết quả đúng (s,t) và bộ giải mã ψ cho kết quả đúng t với xác suất lớn hơn hoặc bằng $(1-\epsilon)$.*

Theo Định nghĩa 2 thì:

$$\Sigma_{x^n \in \mathcal{X}^n} f(x^n|s,t) P^n_{Y|X}\{\varphi(y^n) = (s,t)|x^n\} \geq 1 - \epsilon$$
$$\Sigma_{x^n \in \mathcal{X}^n} f(x^n|s,t) P^n_{Z|X}\{\psi(z^n) = t|x^n\} \geq 1 - \epsilon$$

Mức độ không thể nhận biết của Eve đối với thông báo riêng của Alice gửi cho Bob được xác định bởi độ bất định $H(S|Z^n)$, giá trị này phụ thuộc vào phân bố chung của ST và bộ mã hóa f.

Định nghĩa 3. *Có thể đạt được bộ ba tham số (R_1, R_e, R_0) khi và chỉ khi tồn tại một chuỗi các tập thông báo $\mathcal{S}_n, \mathcal{T}_n$ và bộ mã hóa - giải mã (f_n, φ_n, ψ_n) tạo thành $(n, \epsilon_n) - transmission$ với $\epsilon \to 0$, sao cho:*

$$\lim_{n\to\infty} \frac{1}{n} \log \|\mathcal{S}_n\| = R_1$$
$$\lim_{n\to\infty} \frac{1}{n} \log \|\mathcal{T}_n\| = R_0'$$
$$\lim_{n\to\infty} \frac{1}{n} H(S_n|Z^n) \geq R_e$$

Trong đó, $H(S_n|Z^n)$ được đánh giá dựa theo giả thiết là cặp thông báo ngẫu nhiên S_nT_n là cùng phân bố trên $S_n \times T_n$. Ký hiệu $\mathcal{R}$ là miền chứa các tập giá trị có thể đạt được của bộ ba tham số tốc độ. Nếu $(R_1, R_e, R_0) \in \mathcal{R}$ chúng ta nói rằng R_1 và R_0 là tốc độ truyền tin có thể đạt được của thông báo riêng và thông báo chung với độ mập mờ R_e.

Thay vì quan tâm đến mã hóa nguồn kênh, Imre Csiszár và János Körner quan tâm đến bài toán ghép nối nguồn kênh. Với hai nguồn không nhớ được ký hiệu là $\mathcal{S}$ (nguồn thông báo riêng) và $\mathcal{T}$ (nguồn thông báo chung), theo đó $S_1T_1, S_2T_2, \ldots$ là các cặp biến ngẫu nhiên, độc lập và cùng phân bố (tuy nhiên S_i và T_i không cần độc lập). Gọi $\bar{S}$ và $\bar{T}$ là các biến chung cho hai nguồn tin. Chúng ta giả sử rằng, hệ thống sử dụng bộ mã hóa ngẫu nhiên block-to-block $(k, n) - encoder$ theo như Định nghĩa 1 với độ dài khối là n và các tập thông báo là $\mathcal{S}^k, \mathcal{T}^k$.

Các thông báo ngẫu nhiên có độ dài k khi này được Alice truyền đi là S^k, T^k. Eve không thể nhận biết được nội dung thông báo riêng (S^k) với độ mập mờ

$$\Delta = \frac{1}{k} H(S^k | Z^n).$$

Để đảm bảo truyền tin tin cậy, thì cả hai giá trị tần suất lỗi trung bình là $E\frac{1}{k} d_H(S^k T^k, \varphi(Y^n))$ và $E\frac{1}{k} d_H(T^k, \psi(Z^n))$ phải nhỏ, trong đó, d_H là khoảng cách Hamming.

Định nghĩa 4. *Một cặp nguồn tin $\bar{S}, \bar{T}$ được gọi là $(R, \Delta) - transmission$ trên BCC, trong đó $R > 0, \Delta \geq 0$, khi và chỉ khi với mọi $\epsilon > 0$ tồn tại một bộ mã hóa $(k, n) - encoder f$ và bộ giải mã (φ, ψ) sao cho:*

$$\frac{k}{n} \geq R - \epsilon \tag{4}$$

$$\frac{1}{k} H(S^k | Z^n) \geq \Delta - \epsilon \tag{5}$$

$$E\frac{1}{k} d_H(S^k T^k, \varphi(Y^n)) \leq \epsilon, \; E\frac{1}{k} d_H(T^k, \psi(Z^n)) \leq \epsilon \tag{6}$$

R ở đây được đề cập như là tốc độ ghép nối nguồn kênh (the rate of source-channel matching)

3.2 Các kết quả chính

Định lý 1. *Tập $\mathcal{R}$ chứa bộ ba tham số (R_1, R_e, R_0) là tập lồi đóng (closed convex set) theo đó tồn tại các biến ngẫu nhiên thỏa mãn chuỗi Markov $U \to V \to X \to YZ$ sao cho điều kiện phân phối của Y được cho bởi X là được xác định trên kênh 1 (kênh chính), tương ứng điều kiện phân phối của Z được cho bởi X là được xác định trên kênh 2 (kênh nghe lén) và*

$$0 \leq R_e \leq R_1 \tag{7}$$

$$R_e \leq I(V;Y|U) - I(V;Z|U) \tag{8}$$

$$R_1 + R_0 \leq I(V;Y|U) + \min[I(U;Y), I(U;Z)] \tag{9}$$

$$0 \leq R_0 \leq \min[I(U;Y), I(U;Z)] \tag{10}$$

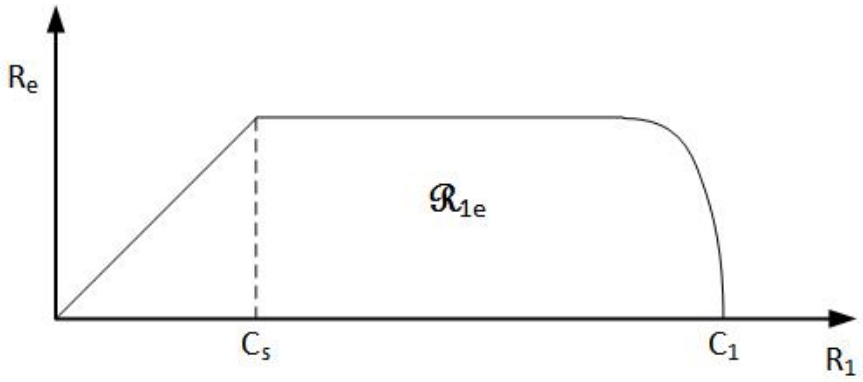
Phần chứng minh các công thức trên được trình bày trong phần IV, V và Phụ lục của tài liệu tham khảo [2].

Định lý 2. *Để cặp nguồn $\bar{S}, \bar{T}$ là $(R, \Delta) - transmissible$ trên kênh BCC (theo Định nghĩa 4) thì điều kiện cần và đủ là:*

$$(RH(\bar{S} \mid \bar{T}), R\Delta, RH(\bar{T}) = (R_1, R_e, R_0) \in \mathcal{R}$$

Phần chứng minh các công thức trên được trình bày trong phần IV và V của tài liệu tham khảo [2].

Khi Eve giải mã thành công thông báo chung do Alice gửi quảng bá cho tất cả mọi người trong hệ thống, thì liệu Eve có thể biết được thông tin gì về thông báo riêng thông qua nội



Hình 5: Miền giá trị $\mathcal{R}_{1e}$ với trường hợp $R_0 = 0$.

dung của thông báo chung. Về mặt lý thuyết thì thông báo chung có thể chứa một phần thông tin nào đó của thông báo riêng. Do đó, cần có điều kiện an toàn là $\Delta \leq H(\bar{S}, \bar{T})$. Thực tế thì điều kiện này đã được chỉ ra trong Định lý 1 (trong (7)) và Định lý 2: là $RH(\bar{S} \mid \bar{T}) \geq R\Delta \;\Leftrightarrow\; H(\bar{S} \mid \bar{T}) \geq \Delta$. Nếu $H(\bar{S} \mid \bar{T}) = \Delta$ thì hai nguồn thông báo là hoàn toàn độc lập và hệ thống đạt tới độ an toàn tuyệt đối, khi này $\Delta = H(\bar{S})$.

Theo Định nghĩa 4, quá trình truyền tin là tuyệt đối an toàn trong tình huống trên khi và chỉ khi tốc độ ghép kênh R thỏa mãn $(RH(\bar{S}), RH(\bar{T}) \in \mathcal{C}_s)$, trong đó, $\mathcal{C}_s$ là miền truyền tin an toàn (the secrecy capacity region) được định nghĩa như sau:

Định nghĩa 5. *Miền truyền tin an toàn $\mathcal{C}_s$ của BCC là một tập các cặp giá trị (R_1, R_0) sao cho $(R_1, R_1, R_0) \in \mathcal{R}$.*

Hệ quả 1: Theo Định nghĩa 5 và Định lý 1, miền $\mathcal{C}_s$ bao gồm các cặp giá trị (R_1, R_0) sao cho tồn tại chuỗi Markov $U \to V \to X \to YZ$ thoả mãn:

$$0 \leq R_1 \leq I(V;Y|U) - I(V;Z|U)$$
$$0 \leq R_0 \leq \min[I(U;Y), I(U;Z)]$$

Xét trường hợp đặc biệt, khi hệ thống không truyền quảng bá thông báo chung ($R_0 = 0$), Ký hiệu miền $\mathcal{R}_{1e}$ chứa các cặp giá trị (R_1, R_e) có thể đạt được, như vậy $(R_1, R_e) \in \mathcal{R}_{1e}$ khi và chỉ khi $(R_1, R_e, 0) \in \mathcal{R}$. Miền giá trị của $\mathcal{R}_{1e}$ được thể hiện như Hình 5.

Theo kết quả của A. D. Wyner ở phần trên, C_s ở đây được định nghĩa là tốc độ truyền thông báo riêng lớn nhất có thể từ Alive tới Bob mà vẫn giữ được bí mật với Eve.

$$C_s \triangleq \max_{(R_1,R_1)\in\mathcal{R}_{1e}} R_1 = \max_{(R_1,0)\in\mathcal{C}_s} R_1 \tag{11}$$

Và trong trường hợp $R_0 = 0$ này, Imre Csiszár và János Körner đã xác định rõ hơn về miền $\mathcal{R}_{1e}$, cũng như các tham số R_1, R_e và C_s thông qua hệ quả sau:

Hệ quả 2: Tập $(R_1, R_e) \in \mathcal{R}_{1e}$ khi và chỉ khi tồn tại chuỗi Markov $U \to V \to X \to YZ$ sao cho $I(U;Y) \leq I(U;Z)$ và $0 \leq R_e \leq I(V;Y|U) - I(V;Z|U)$. Do vậy: $R_e \leq R_1 \leq I(V;Y)$

và hơn nữa $C_s = \max_{V \to X \to YZ}[I(V;Y) - I(V;Z)]$. (Xem chứng minh trong tài liệu tham khảo [2], phần III.)

Với giả thiết kênh truyền từ Alice đến Bob tốt hơn (more capable) kênh truyền từ Alice đến Eve, nghĩa là với mọi đầu vào X thì

$$I(X;Y) \geq I(X;Z). \tag{12}$$

Cũng tương tự, với giả thiết kênh truyền từ Alice đến Bob ít nhiễu hơn (less noisy) kênh truyền từ Alive tới Eve khi truyền thông báo riêng, nghĩa là với mọi $V \to X \to YZ$ thì

$$I(V;Y) \geq I(V;Z). \tag{13}$$

Với các giả thiết trên, Imre Csiszár và János Körner đưa ra tiếp kết quả sau:

Hệ quả 3: Khả năng truyền tin an toàn C_s (secrecy capacity) là luôn dương trừ khi kênh truyền từ Alive đến Eve ít nhiễu hơn kênh truyền từ Alive đến Bob. (Chứng minh: theo công thức trên)

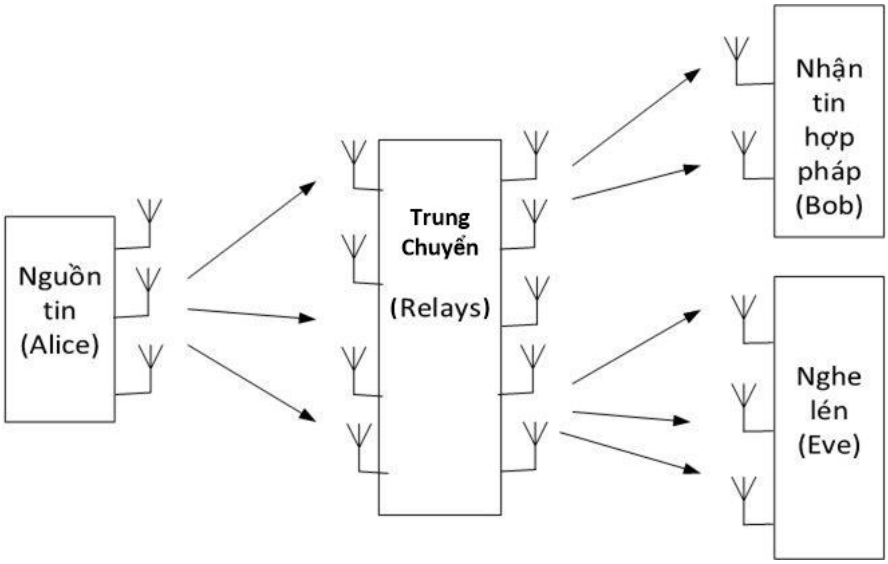
Như vậy (12;13), Imre Csiszár và János Körner đã xem xét mô hình truyền tin đồng thời cả thông báo quảng bá và thông báo riêng cho một người mà vẫn giữ bí mật với người khác trong hệ thống; đã đặc trưng hóa các tham số tốc độ truyền tin theo quan điểm chất lượng thông tin trong lý thuyết truyền tin. Theo đó, về mặt lý thuyết là có thể tính toán được miền chứa các bộ giá trị của các tham số truyền tin để đảm bảo an toàn. Tuy nhiên, tại thời điểm đó, việc tính toán thực tế được xem là rất khó khăn. Các kết quả ban đầu này là rất quan trọng để mở ra một lĩnh vực nghiên cứu mới trong bảo mật thông tin mạng không dây. Tuy nhiên, phải sau khoảng 30 năm công bố, vấn đề này mới thực sự được giới khoa học quan tâm. Phần 4 sau đây sẽ giới thiệu ngắn gọn về tình hình nghiên cứu của vấn đề này trong thời gian gần đây.

4 Các nghiên cứu hiện nay và Kết luận

Từ những kết quả ban đầu sơ khai của A. D. Wyner và Imre Csiszár, János Korner về một khẳng định rằng có thể truyền tin bí mật trong hệ thống truyền tin không dây có sự xuất hiện của người nghe lén, dựa trên lý thuyết truyền tin mà không dùng đến phương pháp mã hóa bảo mật. Trong những năm gần đây, khi kỹ thuật truyền tin không dây đã có nhiều thay đổi, thì ý tưởng này đã được nhiều nhà khoa học trên khắp thế giới tập trung nghiên cứu và ngày càng có tính ứng dụng thực tế cao.

4.1 Tình hình nghiên cứu hiện nay

Kết quả điều tra, tổng hợp và phân tích từ hơn 300 kết quả đã công bố về lĩnh vực này trong [3], trong đó đa phần là những kết quả được công bố trong khoảng từ năm 2008 đến năm 2014, cho thấy: vấn đề bảo mật truyền tin tại tầng vật lý trong mạng không dây đang được xem là hướng đi mới đầy tiềm năng, với các nghiên cứu hiện nay tập trung trên kênh fading có sự hỗ trợ của các trạm trung chuyển (relays). Hay nói cách khác là ý tưởng của A. D.



Hình 6: Mô hình truyền tin Multiple-Input Multiple-Outpute Muiltiple-Eavesdropper

Wyner ngày càng trở lên hiện thực khi có sự hỗ trợ của hai phương pháp truyền tin chủ yếu là Relay và Cooperative. Khi đó vấn đề bảo mật theo lý thuyết thông tin vẫn được đảm bảo ngay cả khi kênh nghe lén tốt hơn kênh chính.

Các bài toán bảo mật truyền tin tầng vật lý được nghiên cứu trong thời gian gần đây tập trung trên các mô hình hệ thống truyền tin sau:

- Dựa theo số lượng ăng ten phát (nguồn phát) và số lượng ăng ten thu của hệ thống truyền tin không dây thì các bài toán bảo mật truyền tin tầng vật lý được nghiên cứu trên các hệ thống như: Hệ thống truyền tin một đầu phát, một đầu thu - SISO (Single-Input Single-Output); Hệ thống truyền tin một đầu phát, nhiều đầu thu - SIMO (Single-Input Multiple-Output); và hệ thống tryền tin nhiều đầu phát, nhiều đầu thu - MIMO (Multiple-Input Multiple-Output).

- Dựa theo số lượng ăng ten thu lén (số lượng người nghe lén) trong hệ thống truyền tin không dây sẽ có các mô hình được nghiên cứu như: Hệ thống nhiều đầu phát, nhiều đầu thu có sự xuất hiện của một đầu nghe lén - MIMOSE (MIMO Single-Eavesdropper) và hệ thống nhiều đầu phát, nhiều đầu thu có sự xuất hiện của nhiều đầu nghe lén - MIMOME (MIMO Multiple-Eavesdropper), như Hình 6.

Các bài toán bảo mật tầng vật lý này cũng được nghiên cứu trên ba mô hình kỹ thuật truyền tin tương tác chính là Decode-and-Forward (DF), Applify-and-Forward (AF) và Co-operative Jamming (CJ).

- Mô hình DF: Hệ thống truyền tin có hỗ trợ bởi Relays hoạt động theo giao thức Decode-and-Forward bao gồm 2 pha chính: Pha thứ nhất, thông báo được truyền từ nguồn phát đến relays. Tại relays thông báo được giải mã nguồn để có được như thông báo ban đầu. Pha thứ hai, tại relays thông báo được mã lại và nhân với hệ số khuếch đại của relays rồi truyền đến nơi nhận. Cả người nhận hợp pháp và người nghe lén đều có thể thu được tín hiệu từ relays. Với công nghệ truyền tin beamforming ([3]), tín hiệu beamforming

có định hướng từ nhiều relays sẽ tương tác với nhau, cộng với tín hiệu nhiễu tại mỗi kênh tương ứng sao cho tại người nhận hợp pháp tín hiệu có thể giải mã và khôi phục về nguồn tin một cách chính xác, còn tại người nghe lén, tín hiệu bị triệt tiêu hoặc ở mức rất thấp và người nghe lén không thể khôi phục được tín hiệu.

- Mô hình AF: Tương tự như mô hình DF, mô hình truyền tin AF cũng có 2 pha chính. Pha 1, tín hiệu được phát từ nguồn đến relays, nhưng ở đây relays không giải mã nguồn tín hiệu như mô hình DF mà để nguyên tín hiệu thu được rồi nhân với hệ số khuếch đại của relays, sau đó truyền tín hiệu đó đến nơi nhận tại pha 2. Do sự tương tác giữa các tín hiệu từ các relays cộng với nhiễu tại mỗi kênh tương ứng mà tín hiệu tại người nhận hợp pháp có thể khôi phục và giải mã thành công, còn tín hiệu tại người nghe lén bị triệt tiêu hoặc rất thấp, nên không thể khôi phục cũng như giải mã.

- Mô hình CJ: hay còn được gọi là Artificial noise design, cũng được nhiều nhà nghiên cứu quan tâm. Trong mô hình này, các relays (còn được gọi là các friendly jammers) sẽ phát các tín hiệu nhiễu có chủ đích để kết hợp với tín hiệu phát từ nguồn đến người nhận và người nghe lén. Các tín hiệu nhiễu từ trạm phát nhiễu chủ động (friendly jammers) sẽ tác động không đáng kể đến tín hiệu từ nguồn truyền đến người nhận hợp pháp, do đó người nhận hợp pháp sẽ khôi phục và giải mã thành công. Còn tại người nghe lén, tín hiệu jamming sẽ triệt tiêu hoàn toàn hoặc phần lớn tín hiệu từ nguồn phát, làm cho người nghe lén không thể khôi phục cũng như giải mã tín hiệu.

Tất cả các bài toán bảo mật truyền tin tầng vật lý ở trên đều đưa về các dạng bài toán tối ưu theo lý thuyết thông tin (information theory), với hai dạng hàm mục tiêu (objective funtion) và ràng buộc (constraint) chủ yếu là: (i) Tối đa hóa tốc độ truyền tin an toàn (secrecy rate) với ràng buộc về công suất phát (power). (ii) Tối thiểu hóa công suất phát với ràng buộc về tốc độ truyền tin an toàn.

4.2 Kết luận

Trong nhiều năm gần đây, khi nói đến bảo mật dữ liệu truyền tin người ta thường nghĩ ngay tới việc ứng dụng các thuật toán mật mã ở tầng trên, với các thuật ngữ thông dụng như hệ mật khóa đối xứng (khóa bí mật), hệ mật khóa bất đối xứng (khóa công khai); các thuật toán mật mã như 3DES, AES, GOST,... bên cạnh đó là các thuật toán sinh khóa và trao đổi khóa (thỏa thuận khóa). Mặc dù hiện nay phương pháp mã hóa bảo mật ở trên vẫn đang đáp ứng tốt nhu cầu bảo mật thông tin, tuy nhiên phương pháp truyền thống này luôn chứa đựng những rủi ro nhất định, vì độ bảo mật được dựa trên độ phức tạp tính toán của bài toán giải mã khi không có khóa hay độ phức tạp tính toán của các phương pháp tấn công. Đặc biệt, khi máy tính lượng tử thực sự được đưa vào ứng dụng thì các căn cứ an toàn trên sẽ bị phá vỡ.

Do vậy, giải quyết bài toán bảo mật truyền tin tại tầng vật lý là hướng đi cần được quan tâm trong thời điểm hiện nay cũng như trong thời gian tới. Bảo mật truyền tin tầng vật lý cho mạng không dây có thể được ứng dụng trong một hệ thống truyền tin an toàn độc lập mà không cần có giải pháp mật mã truyền thống, khi đó các thông báo bí mật sẽ được truyền với tốc độ bảo mật (secrecy rate) là C_s. Bảo mật tầng vật lý cũng có thể được dùng kết hợp với các phương pháp bảo mật bằng mật mã truyền thống, khi đó bảo mật tầng vật lý có thể

sử dụng để truyền khóa bí mật cho các thuật toán mật mã. Hoặc có thể cài đặt song song cả bảo mật tầng vật lý với mật mã truyền thống để tăng mức độ bảo mật.

Tài liệu

[1] A.D. Wyner, *The Wire-tap channel*, Bell Sys. Tech. Journal, Vol 54, 1975.

[2] IMRE CSIS and JALNOS KORNER, *The Broadcast Channels With Confidential Messages*, IEEE transaction on Information Thoery, Vol IT-24, No. 3, May 1978.

[3] A. Mukherjee and S. A. A. Fakoorian and J. Huang and A. L. Swindlehurst *Principles of Physical Layer Security in Multiuser Wireless Networks: A Survey*, IEEE Communications Surveys Tutorials, Vol 16, 2014.