

MÔ HÌNH DỰ ĐOÁN CHO MÃ ĐỘC TRONG MẠNG CẢM BIẾN KHÔNG DÂY

Trần Hải Anh, Đỗ Xuân Chợt, Nguyễn Đức Việt
Học Viện Công Nghệ Bưu Chính Viễn Thông

Tóm tắt: Mô hình truyền nhiễm là một dạng mô hình toán học đã và đang được nghiên cứu trong nhiều lĩnh vực khác nhau, đặc biệt kể đến như dịch tễ học, thống kê ... Dựa trên sự tương đồng giữa virus sinh học và mã độc, bài viết này giới thiệu một mô hình được mô tả bằng hệ phương trình vi phân thông thường. Nó dự đoán sự lây lan của mã độc dựa trên bản chất của mã độc đó. Đối tượng của chúng có thể là hướng tới các doanh nghiệp hay là trên các máy tính cá nhân của người dùng thông thường. So với mô hình dịch bệnh đơn giản, chúng tôi đề xuất một mô hình mới với các nút có mối liên hệ mật thiết thông qua việc lây truyền mã độc cho nhau. Từ đó cho phép chúng ta có thể nhanh chóng đưa ra các chính sách nhằm kiểm soát sự lây lan một cách phù hợp.

Từ khóa: mạng cảm biến không dây, mã độc, phương trình đặc trưng, hệ phương trình vi phân.

I. GIỚI THIỆU

Mạng cảm biến không dây (WSN) là một cấu trúc, là sự kết hợp các khả năng cảm biến, xử lý thông tin và các thành phần liên lạc để tạo khả năng quan sát, phân tích và phản ứng lại với các sự kiện và hiện tượng xảy ra trong môi trường cụ thể nào đó. Vài thập kỷ qua, với sự phát triển mạnh mẽ của mạng cảm biến không dây đã khiến nhiệm vụ đảm bảo an toàn thông tin trở nên đầy thách thức trước các cuộc tấn công bằng mã độc gây ra. Đặc điểm lan truyền mã độc trong WSN tương tự như cách mà một căn bệnh truyền nhiễm lây từ người sang người. Mô hình toán học đầu tiên SIR, mô tả sự lan truyền này được giới thiệu bởi Kermack và McKendrick [1]. Trong mô hình toán học dạng ngăn được mô tả qua hệ phương trình (1), dân số được chia thành 3 nhóm khác nhau bao gồm: S (Mẫn cảm) đại diện cho người khỏe mạnh và có thể mắc bệnh, I (Lây nhiễm) đại diện cho người đang nhiễm bệnh và có thể lây cho người khác, R (Hồi phục) đại diện cho những người đã khỏi bệnh và không có khả năng mắc bệnh lại và sau đó các tác giả phân tích đặc trưng dịch tễ và dự báo chiều hướng bùng phát của dịch bệnh.

Tác giả liên hệ: Trần Hải Anh,

Email: haianh121293@gmail.com

Đến tòa soạn: 22/4/2024, chỉnh sửa: 24/5/2024, chấp nhận đăng: 12/6/2024.

$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta IS}{N} \\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I\end{aligned}\quad (1)$$

Lây cảm hứng từ sự tương đồng giữa sự lan truyền dịch bệnh trong quần thể sinh học và sự lan truyền mã độc trên không gian mạng, mô hình này đã được áp dụng để nghiên cứu mô hình lan truyền mã độc bởi Youssef và cộng sự [2], Rey và cộng sự [3]. Hơn nữa, trong những năm gần đây, nhiều mô hình tuyến nhiễm mã độc dựa trên mô hình SIR cổ điển đã được phát triển. Ví dụ như Prajapati [4] đưa ra mô hình mô tả sự lây lan mã độc thông qua thiết bị di động, Hosseini và cộng sự [5] đề xuất mô hình SEIRS-C để mô tả quá trình tin đồn lan truyền trên mạng xã hội hay Liu và cộng sự [6] đề xuất mô hình có hai loại mã độc cạnh tranh lẫn nhau trên mạng.

Phần lớn các nghiên cứu trên đều chủ yếu tập trung vào mô hình hóa sự lan truyền của mã độc trên không gian mạng mà bỏ qua bản chất thực tế của chúng. Để có một chiến lược kiểm soát và ngăn chặn sự lây lan của mã độc trên mạng cảm biến không dây, một mô hình truyền nhiễm được giới thiệu trong bài viết này là mô hình SIR kết hợp. Tính mới của nghiên cứu này là khám phá hành vi của mã độc tác động đến quá trình lây lan của chúng trên WSN thông qua mô hình toán học dạng ngăn. Phần còn lại của bài viết được tổ chức như sau: Phần 2 sẽ trình bày mục các nghiên cứu liên quan. Tiếp đó là mục xây dựng kiến trúc mô hình đề xuất tại phần 3. Quá trình phân tích toán học thực hiện ở phần 4. Thử nghiệm mô phỏng được minh họa trong phần 5. Cuối cùng là kết luận được trình bày tại phần 6.

II. CÁC NGHIÊN CỨU LIÊN QUAN

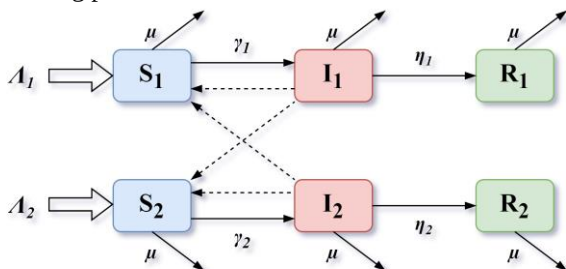
Trong nghiên cứu của Pastor-Satorras và Vespignani [9] được biết đến như một công trình tiên phong trong lĩnh vực nghiên cứu các mô hình toán học có thể mô tả được sự lây nhiễm của các đối tượng độc hại trên mạng phức hợp. Cụ thể hơn, công trình nghiên cứu này cho ra mô hình ngăn SIS dựa trên mạng và trình bày một nghiên cứu chi tiết về đặc tính dịch tễ cơ bản và kết quả giải số cho mô hình dịch tễ đã đề xuất. Một đề xuất khác bởi Huo và cộng sự [10] đã cho ra một mô hình dịch ba ngăn với các trạng thái mẫn cảm, lan truyền virus và phục hồi để mô tả quá trình lan truyền virus trên mạng quy mô tự do. Trong nghiên cứu

này, các tác giả xác định hệ số lây nhiễm thứ phát R_0 và thảo luận vai trò của ngưỡng giá trị của R_0 đối với đáng điều kiện cận của mô hình. Trong nghiên cứu của Li và Yousef [11] đã giới thiệu mô hình dịch SIRS dựa trên mô hình mạng kết hợp với hàm xử lý mã độc bão hòa để mô tả rõ hơn kịch bản thực tế khi số lượng người nhiễm bệnh có thể vượt quá khả năng điều trị. Bài báo này đã chỉ ra hệ số lây nhiễm cơ bản R_0 và xác định sự ổn định tiệm cận của điểm cân bằng cùng với phân tích hiện tượng rẽ nhánh lồi tại $R_0 = 1$. Một điểm mới của nghiên cứu này là tác giả đã sử dụng hàm xử lý mã độc bão hòa thay vì tuyến tính để phù hợp hơn với kịch bản bùng phát dịch trong thực tế. Một yếu tố quan trọng không kém là thời gian đã được Wang và cộng sự [12] đưa ra nhằm tối ưu chiến lược tiêm chủng và chứng minh được rằng việc đưa yếu tố thời gian vào mô hình đã giúp kiểm soát sự lây lan của dịch bệnh một cách hiệu quả, họ cũng áp dụng Pontryagin's maximum principle cho chiến lược kiểm soát tối ưu. Một nghiên cứu khác do Zheng và cộng sự [13] đề xuất mô hình SLBQR cho bài toán biểu diễn sự lây lan của worm trên internet di động có tính đến độ trễ và tốc độ cách ly nhóm nhiễm bệnh, các tác giả đã chứng minh rằng không có chiến lược tiêm chủng nào đạt được khả năng miễn dịch vĩnh viễn và lợi ích của các chiến lược cách ly. Đề xuất sử dụng mô hình Bayesian Networks để chi phối động lực lan truyền của mã độc đã được đưa ra bởi Guillen và cộng sự [14], mô hình này xét đến việc đối phó với các thiết bị lây nhiễm trong khu vực lân cận. Từ những nghiên cứu trên chúng tôi thấy được cần có những nghiên cứu để giải quyết bài toán dự đoán sự lây lan của mã độc dựa trên hành vi tấn công của mã độc.

III. XÂY DỰNG MÔ HÌNH

Trong phần này, một mô hình truyền nhiễm mới được xây dựng để có khả năng mô tả chính xác hành vi lây lan mã độc dựa trên mô hình SIR nổi tiếng đã được chúng tôi đề cập tại mục I. Tất cả các nút trong WSN đều xác định thành hai loại với ba trạng thái (Mẫn cảm - Lây nhiễm - Phục hồi):

- $S_1(t)$ (rate of susceptible nodes): Tỷ lệ nhóm máy không quan trọng dễ lây nhiễm mã độc tại thời điểm t .
- $I_1(t)$ (rate of infectious nodes): Tỷ lệ nhóm máy không quan trọng bị lây nhiễm mã độc tại thời điểm t .
- $R_1(t)$ (rate of recovered nodes): Tỷ lệ nhóm máy không quan trọng phục hồi tại thời điểm t .
- $S_2(t)$ (rate of susceptible nodes): Tỷ lệ nhóm máy quan trọng dễ bị lây nhiễm mã độc tại thời điểm t .
- $I_2(t)$ (rate of infectious nodes): Tỷ lệ nhóm máy quan trọng bị nhiễm mã độc tại thời điểm t .
- $R_2(t)$ (rate of recovered nodes): Tỷ lệ nhóm máy quan trọng phục hồi tại thời điểm t .



Hình 1: Biểu đồ dòng của hệ thống

Bảng 1: Định nghĩa các tham số

Ký hiệu	Ý nghĩa
Λ_1	Hệ số gia tăng vào S_1
Λ_2	Hệ số gia tăng vào S_2
γ_1	Hệ số chuyển đổi từ S_1 tới I_1
γ_2	Hệ số chuyển đổi từ S_2 tới I_2
η_1	Hệ số phục hồi từ I_1 tới R_1
η_2	Hệ số phục hồi từ I_2 tới R_2
μ	Tỷ lệ sụt giảm do hết năng lượng

Xây dựng mô hình đề xuất được biểu diễn trong hình 1 với các giả định về cơ chế lan truyền mã độc trong WSN:

(i) Đầu tiên chúng tôi cho rằng tất cả các máy tính trong cùng một nhóm đều được xây dựng đồng nhất về mọi mặt như hệ điều hành, ứng dụng, tường lửa, hệ thống bảo mật ...v.v.

(ii) Xây dựng đặc tính cho các nút nhạy cảm thuộc nhóm không quan trọng. Nhóm S_1 biểu diễn trạng thái các máy tính chưa nhiễm mã độc nhưng lại rất nhạy cảm với các mã độc. Những nút này sẽ bị lây nhiễm bởi các nút nhiễm mã độc từ cả hai nhóm với tốc độ γ_1 thông qua giao thức truyền thông, các nút thay đổi trạng thái từ nhóm S_1 sang nhóm I_1 . Các nút được bổ sung vào hệ thống sẽ tham gia vào nhóm S_1 với tốc độ Λ_1 và đồng thời một số nút bị loại khỏi nhóm S_1 do hết năng lượng với tốc độ μ . Vì vậy mối quan hệ chuyển giao giữa S_1 và các nhóm khác có thể biểu diễn như sau:

$$S_1' = \Lambda_1 - (\gamma_1 + \mu)S_1 \quad (2)$$

(iii) Xây dựng đặc tính cho những nút bị nhiễm mã độc thuộc nhóm máy không quan trọng. Nhóm I_1 chứa các máy bị nhiễm mã độc và có khả năng lây nhiễm mã độc sang nhóm các nút nhạy cảm. Mật độ của các nút nhiễm mã độc thuộc nhóm I_1 sẽ tăng lên khi các nút dễ lây nhiễm S_1 bị nhiễm mã độc. Hệ số η_1 đại diện cho tốc độ chuyển đổi một số nút từ nhóm I_1 sang nhóm phục hồi R_1 sau khi mã độc đã được loại bỏ. Một số nút ở nhóm I_1 sẽ bị loại bỏ do hệ thống hệ thống cập nhật trạng thái và loại bỏ đi những nút cảm biến trong WSN đã hết năng lượng với tốc độ là μ . Phương trình (3) dưới đây sẽ cho thấy mối quan hệ chuyển tiếp giữa nhóm I_1 và các nhóm khác:

$$I_1' = \gamma_1 S_1 - (\eta_1 + \mu)I_1 \quad (3)$$

(iv) Xây dựng đặc tính cho các nút tại trạng thái phục hồi thuộc nhóm máy không quan trọng. Nhóm R_1 bao gồm các máy đã loại bỏ mã độc và chúng tôi cũng giả định thêm rằng máy không có khả năng bị tái nhiễm. Gọi tốc độ các nút nhiễm mã độc I_1 sau đó chúng được hồi phục và đưa vào nhóm R_1 là η_1 . Chúng tôi giả định rằng tại mỗi thời điểm hệ thống cập nhật, có một lượng nhất định nút ở nhóm R_1 sẽ bị loại khỏi hệ thống do cạn năng lượng với tốc độ μ . Do đó mối quan hệ chuyển giao giữa R_1 và I_1 có thể biểu diễn như sau:

$$R_1' = \eta_1 I_1 - \mu R_1 \quad (4)$$

Có thể nhận thấy các giả định có phần thiếu thực tế như việc coi các máy tính là hoàn toàn đồng nhất, tuy nhiên điều này lại giúp đơn giản hóa mô hình toán học để chúng tôi có thể dễ dàng thực hiện nghiên cứu hành vi lây nhiễm của mã độc. Tiếp tục, đối với mô hình SIR còn lại cho nhóm máy quan trọng chúng ta cũng thực hiện việc xây dựng các phương trình một cách tương tự. Từ các giả định (i), (ii), (iii) và (iv) chúng tôi thu được mô hình SIR kết hợp với khả năng kiểm soát tốc độ lan truyền mã độc cho bởi tập các phương trình vi phân như sau:

$$\begin{aligned} S'_1 &= \Lambda_1 - (\gamma_1 + \mu)S_1 \\ I'_1 &= \gamma_1 S_1 - (\eta_1 + \mu)I_1 \\ R'_1 &= \eta_1 I_1 - \mu R_1 \\ S'_2 &= \Lambda_2 - (\gamma_2 + \mu)S_2 \\ I'_2 &= \gamma_2 S_2 - (\eta_2 + \mu)I_2 \\ R'_2 &= \eta_2 I_2 - \mu R_2 \end{aligned} \quad (5)$$

IV. PHÂN TÍCH MÔ HÌNH

A. Phân tích các hệ số

Để làm rõ hơn phương trình vi phân (5) chúng tôi sẽ sử dụng các loại tỷ lệ nhằm mô tả mối quan hệ giữa các tham số trong hai nhóm máy. Coi M là tổng số máy tính trong mạng cảm biến không dây tại thời điểm t bất kỳ với $\forall t > 0$. Với đầu vào của hai nhóm là Λ_1 và Λ_2 , ta có tỷ lệ gia tăng lượng máy tính vào mô hình là: $\Lambda = \Lambda_1 + \Lambda_2$. Dựa theo luật phân phối nhị thức, chúng tôi coi xác suất Λ_1 là p , suy ra xác suất Λ_2 là $1 - p$.

Tiếp theo cần định nghĩa tham số mô tả tổng số máy tính trong nhóm thứ nhất là m_1 tại thời điểm t được xác định bởi $S_1(t) + I_1(t) + R_1(t)$ và tổng số máy tính trong nhóm hai là m_2 tại thời điểm t được tính bằng $S_2(t) + I_2(t) + R_2(t)$. Mô hình cũng biểu diễn sự suy giảm do nút cảm biến hết năng lượng sau mỗi lần hệ thống cập nhật với tốc độ μ trên tất cả các trạng thái nên sự thay đổi số máy tính tại thời điểm t được xác định như sau:

Đối với nhóm máy không quan trọng:

$$m'_1 = p\Lambda - \mu m_1$$

Đối với nhóm máy quan trọng:

$$m'_2 = \Lambda(1 - p) - \mu m_2$$

Từ hai phương trình trên ta thu được mức độ thay đổi tổng thể của cả mô hình:

$$M' = \Lambda - \mu(M_1 + M_2) = \Lambda - \mu M$$

Do sự bùng phát lây nhiễm mã độc trên không gian mạng diễn ra với tốc độ nhanh hơn so với các virus lây nhiễm trên cơ thể người như nhóm bệnh cúm. Vì vậy chúng tôi đã giả sử rằng tổng số máy tính trong WSN là không đổi tại mọi thời điểm: $\Lambda - \mu M = 0$ với $M = S_1 + I_1 + R_1 + S_2 + I_2 + R_2 = 1$, suy ra sự gia tăng máy tính trên toàn bộ hệ thống mạng là: $\Lambda = \mu M = \mu$.

Chúng tôi gọi k là hệ số tiếp xúc của nhóm máy quan trọng nhằm biểu diễn sự tác động qua lại giữa nhóm **I** là

các nút nhiễm mã độc và nhóm **S** đại diện cho các nút dễ lây nhiễm. Hơn nữa ta coi tỉ lệ tiếp xúc từ các nhóm **I** sang nhóm **S**₁ là tương đương nhau, tương tự như vậy với các nhóm **I** sang nhóm **S**₂ đều có tỉ lệ chuyển đổi là bằng nhau là $k(I_1 + I_2)$, tiếp đó chỉ cần xét tới sự khác biệt giữa γ_1 và γ_2 . Sau khi định nghĩa hệ số k , chúng tôi tiếp tục biểu diễn tỉ lệ chuyển đổi của nhóm máy không quan trọng so với nhóm máy quan trọng bằng hệ số α được tính bằng $\frac{\gamma_1}{\gamma_2}$, nhờ đó mô hình có được khả năng mô tả mối quan hệ giữa hai hệ số γ_1 và γ_2 . Kết hợp lại ta có biểu thức: $\gamma_1 = \alpha\gamma_2 = \alpha k(I_1 + I_2)$ là phương trình biểu diễn mối quan hệ giữa hệ số chuyển đổi của cả hai nhóm máy với hệ số tiếp xúc.

Thực hiện giống như hệ số α , chúng tôi cũng viết lại tốc độ phục hồi cho hai nhóm máy là η_1 và η_2 bằng việc mô tả sự tương quan giữa hai tham số này thông qua hệ số β để chỉ tỉ lệ phục hồi của nhóm máy không quan trọng so với nhóm máy quan trọng: $\eta_1 = \beta\eta_2$.

Thông qua việc phân tích các tham số chúng tôi xây dựng lại phương trình vi phân (5) bằng một phương trình mới có khả năng mô tả được mối liên hệ giữa hai nhóm máy:

$$\begin{aligned} S'_1 &= p\mu - (\alpha\gamma_2 + \mu)S_1 \\ I'_1 &= \alpha\gamma_2 S_1 - (\beta\eta_2 + \mu)I_1 \\ R'_1 &= \beta\eta_2 I_1 - \mu R_1 \\ S'_2 &= \mu - p\mu - (\gamma_2 + \mu)S_2 \\ I'_2 &= \gamma_2 S_2 - (\eta_2 + \mu)I_2 \\ R'_2 &= \eta_2 I_2 - \mu R_2 \end{aligned} \quad (6)$$

B. Xác định tọa độ điểm cân bằng

Tọa độ của các điểm cân bằng trong mô hình đề xuất của chúng tôi chính là nghiệm của hệ ứng với vế phải của hệ phương trình (6) bằng 0.

$$\begin{cases} S_1 + S_2 + I_1 + I_2 + R_1 + R_2 = 1 \\ p\mu - (\alpha\gamma_2 + \mu)S_1 = 0 \\ \alpha\gamma_2 S_1 - (\beta\eta_2 + \mu)I_1 = 0 \\ \beta\eta_2 I_1 - \mu R_1 = 0 \\ \mu - p\mu - (\gamma_2 + \mu)S_2 = 0 \\ \gamma_2 S_2 - (\eta_2 + \mu)I_2 = 0 \\ \eta_2 I_2 - \mu R_2 = 0 \end{cases} \quad (7)$$

Để có thể xác định được các điểm cân bằng tự do $E_f(S_1^f, S_2^f, I_1^f, I_2^f, R_1^f, R_2^f)$ và điểm cân bằng bệnh $E_e(S_1^e, S_2^e, I_1^e, I_2^e, R_1^e, R_2^e)$, chúng tôi thực hiện giải hệ phương trình (7) và thu được hai điểm cân bằng:

Điểm cân bằng tự do:

$$E_f = (p, 1 - p, 0, 0, 0, 0) \quad (8)$$

Điểm cân bằng bệnh:

$$E_e = \begin{pmatrix} \frac{p\mu}{\alpha\gamma_2 + \mu} \\ \frac{(1-p)\mu}{\gamma_2 + \mu} \\ \frac{\alpha\gamma_2 p\mu}{(\alpha\gamma_2 + \mu)(\beta\eta_2 + \mu)} \\ \frac{\gamma_2(1-p)\mu}{(\gamma_2 + \mu)(\eta_2 + \mu)} \\ \frac{\beta\eta_2 \alpha\gamma_2 p}{(\alpha\gamma_2 + \mu)(\beta\eta_2 + \mu)} \\ \frac{\eta_2 \gamma_2 (1-p)}{(\gamma_2 + \mu)(\eta_2 + \mu)} \end{pmatrix}^T \quad (9)$$

C. Xác định hệ số lây nhiễm

Theo dịch tễ học, hệ số lây nhiễm cơ bản R_0 là số trường hợp nhiễm bệnh mới dự kiến được tạo ra trực tiếp bởi một trường hợp nhiễm bệnh ban đầu. Ngưỡng này cho biết sự lan truyền tồn tại hay không tồn tại trong cộng đồng với giá trị R_0 càng lớn thì dịch càng khó kiểm soát, điều này cũng được áp dụng vào mô hình dự đoán lây lan mã độc trong không gian mạng cảm biến không dây. Đây là thông số chính cho việc biểu diễn số lượng máy tính trung bình bị nhiễm mã độc trong một giai đoạn.

Để xác định hệ số lây nhiễm cơ bản R_0 chúng ta có thể tính được thông qua các mô hình toán học. Một số phương pháp thường được sử dụng trong dịch tễ học như: Ước lượng hợp lý cực đại (Maximum likelihood estimation); Tăng trưởng theo cấp số nhân (Exponential growth). Tuy nhiên ở đây chúng tôi đã dựa trên ma trận thế hệ tiếp theo (next-generation matrix) được đề xuất bởi Diekmann và cộng sự [7] kết hợp với nghiên cứu của Driessche và Watmough [8]. Ưu điểm chính của ma trận thế hệ tiếp theo là cho phép nhà nghiên cứu bỏ qua các nhóm không bị nhiễm mã độc mà chỉ tập trung vào các nhóm bị nhiễm. Trong mô hình được đề xuất đã biểu diễn hai nhóm nút mạng trạng thái bị nhiễm mã độc nên chúng tôi sẽ định nghĩa: $\mathbf{u} = (I_1, I_2)^T$, sau đó thu được hệ con với khả năng tái nhiễm mã độc được lược bỏ:

$$\mathbf{u}' = (D_u \mathcal{F} + D_u \mathcal{V}) \mathbf{u} \quad (10)$$

Trước khi xác định ma trận $D_u \mathcal{F}$ và ma trận $D_u \mathcal{V}$, chúng tôi lập phương trình vi phân (11):

$$\mathbf{u}' = \mathcal{F} + \mathcal{V} \quad (11)$$

với:

$$\mathcal{F} = \begin{pmatrix} \alpha k S_1 (I_1 + I_2) \\ k S_2 (I_1 + I_2) \end{pmatrix} \quad \text{và} \quad \mathcal{V} = \begin{pmatrix} -I_1 (\beta \eta_2 + \mu) \\ -I_2 (\eta_2 + \mu) \end{pmatrix}$$

Chúng tôi gọi $D_u \mathcal{F}$ và $D_u \mathcal{V}$ lần lượt là ma trận của tất cả các đạo hàm riêng cấp một của $\mathcal{F}$ và $\mathcal{V}$ được tính tại điểm cân bằng tự do P_f^* không có máy tính nhiễm mã độc trong hệ thống:

$$D_u \mathcal{F} = \begin{pmatrix} \alpha k p & \alpha k p \\ k(1-p) & k(1-p) \end{pmatrix} \quad (12)$$

$$D_u \mathcal{V} = \begin{pmatrix} -\beta \eta_2 - \mu & 0 \\ 0 & -\eta_2 - \mu \end{pmatrix} \quad (13)$$

Sau khi xác định được hai ma trận $D_u \mathcal{F}$ và $D_u \mathcal{V}$ được biểu diễn qua phương trình (12) và (13), chúng tôi tiếp tục tính ma trận nghịch đảo của $D_u \mathcal{V}$, sau đó tính tích hai ma trận $-D_u \mathcal{F}$ và $(D_u \mathcal{V})^{-1}$ ta thu được ma trận thế hệ tiếp theo với miền lớn:

$$\begin{pmatrix} \frac{\alpha k p}{\beta \eta_2 + \mu} & \frac{\alpha k p}{\eta_2 + \mu} \\ \frac{k(1-p)}{\beta \eta_2 + \mu} & \frac{k(1-p)}{\eta_2 + \mu} \end{pmatrix} \quad (14)$$

Cuối cùng, chúng tôi tính được hệ số lây nhiễm cơ bản R_0 thông qua ma trận (14) thu được ở trên:

$$R_0 = \frac{k(1-p)}{\eta_2 + \mu} + \frac{\alpha k p}{\beta \eta_2 + \mu} \quad (15)$$

Theo định nghĩa về hệ số lây nhiễm cơ bản đã trình bày ở trên: Khi hệ số lây nhiễm cơ bản R_0 nhỏ hơn 1 nghĩa là mô hình đạt trạng thái cân bằng tự do và khi hệ số lây nhiễm cơ bản R_0 lớn hơn 1 mô hình luôn có mã độc và ổn định tại điểm cân bằng bệnh.

D. Phân tích tính ổn định

1) Tính ổn định của điểm cân bằng tự do

Dựa trên điểm cân bằng tự do E_f thu được trong mục IV-B. Chúng tôi xét tính ổn định địa phương của điểm cân bằng tự do E_f được tìm bằng phương pháp xét dấu các giá trị riêng của ma trận Jacobi tại E_f . Xét ma trận Jacobi ($J_{S_1^f, R_1^f, S_2^f, R_2^f}$) với $S_1^f = p$, $S_2^f = 1-p$, $R_1^f = R_2^f = 0$:

$$\begin{pmatrix} -\mu & 0 & 0 & 0 \\ 0 & -\mu & 0 & 0 \\ 0 & 0 & -\mu & 0 \\ 0 & 0 & 0 & -\mu \end{pmatrix} \quad (16)$$

Từ ma trận (16) chúng tôi tìm được các giá trị riêng của ma trận này là: $\lambda_1 = -\mu$, $\lambda_2 = -\mu$, $\lambda_3 = -\mu$ và $\lambda_4 = -\mu$. Có thể thấy mọi giá trị riêng $\lambda_1, \lambda_2, \lambda_3$ và λ_4 luôn âm vì tham số μ luôn dương. Suy ra tính ổn định địa phương của mô hình tại điểm E_f phụ thuộc vào ma trận Jacobi ($J_{I_1^f, I_2^f}$) sau đây:

$$\begin{pmatrix} \alpha k p - \beta \eta_2 - \mu & \alpha k p \\ k(1-p) & k(1-p) - \eta_2 - \mu \end{pmatrix} \quad (17)$$

Chúng tôi xác định vết và định thức của $J_{I_1^f, I_2^f}$:

$$\text{tr}(J) = k - kp + \alpha kp - \beta \eta_2 - \eta_2 - 2\mu$$

$$\det(J) = [\beta \eta_2^2 + \beta \eta_2 \mu + \eta_2 \mu + \mu^2](1 - R_0)$$

Như đã biết, nếu định thức dương thì giá trị riêng sẽ cùng dấu với vết. Ta thấy rằng định thức luôn dương nếu $R_0 < 1$, từ đó suy ra $\beta \eta_2 + \mu > \alpha kp$ và $\eta_2 + \mu > k - kp$. Áp dụng hai bất đẳng thức này vào vết của $J_{I_1^f, I_2^f}$, dễ dàng nhận thấy vết của $J_{I_1^f, I_2^f}$ luôn nhỏ hơn 0 nếu mọi tham số thỏa mãn điều kiện hệ số lây nhiễm cơ bản nhỏ hơn 1. Từ đây chúng tôi có thể kết luận rằng giá trị riêng của ma trận tạo bởi I_1, I_2 luôn âm.

Khi hệ số lây nhiễm cơ bản $R_0 < 1$ thì mô hình đề xuất của chúng tôi ổn định tiệm cận tại điểm cân bằng tự do $E_f(p, 1-p, 0, 0, 0)$.

2) Tính ổn định của điểm cân bằng bệnh

Sau khi đánh giá tính ổn định địa phương tại điểm cân bằng tự do, chúng tôi tiếp tục xét đến tính ổn định của mô hình tại điểm cân bằng bệnh E_e . Để kiểm tra tính ổn định

địa phương của điểm cân bằng bệnh chúng tôi sử dụng tiêu chuẩn ổn định Routh-Hurwitz [15], tiêu chuẩn ổn định này được đề xuất một cách độc lập bởi hai nhà toán học Edward John Routh [16] và Adolf Hurwitz [17] vào cuối thế kỷ 19.

Đây là một phương pháp đơn giản nhưng mang lại hiệu quả mạnh mẽ so với nhiều phương pháp khác đòi hỏi tính toán phức tạp, giúp ta không cần tìm nghiệm của phương trình đặc trưng bậc cao mà chỉ cần sắp xếp các hệ số trong đa thức đặc trưng vào bảng Routh và thực hiện xác định xem hệ thống có ổn định hay không. Ma trận Jacobi của hệ phương trình (6) tại điểm cân bằng bệnh E_e đã bỏ qua giai đoạn phục hồi như sau:

$$\begin{pmatrix} -\alpha\gamma_2 - \mu & \frac{-\alpha k p \mu}{\alpha\gamma_2 + \mu} & 0 & \frac{-\alpha k p \mu}{\alpha\gamma_2 + \mu} \\ \alpha\gamma_2 & \frac{\alpha k p \mu}{\alpha\gamma_2 + \mu} - \beta\eta_2 - \mu & 0 & \frac{\alpha k p \mu}{\alpha\gamma_2 + \mu} \\ 0 & \frac{k p \mu - k \mu}{\gamma_2 + \mu} & -\gamma_2 - \mu & \frac{k p \mu - k \mu}{\gamma_2 + \mu} \\ 0 & \frac{k \mu - k p \mu}{\gamma_2 + \mu} & \gamma_2 & \frac{k \mu - k p \mu}{\gamma_2 + \mu} - \eta_2 - \mu \end{pmatrix} \quad (18)$$

Các giá trị riêng nhận được thông qua việc giải phương trình đặc trưng có dạng:

$$C_1\lambda^4 + C_2\lambda^3 + C_3\lambda^2 + C_4\lambda + C_5 = 0 \quad (19)$$

với:

$$C_1 = 1$$

$$C_2 = 4\mu - \frac{k\mu}{\mu + \gamma_2} + \frac{k p \mu}{\mu + \gamma_2} - \frac{\alpha k p \mu}{\mu + \alpha\gamma_2} + \beta\eta_2 + \eta_2 + \alpha\gamma_2 + \gamma_2$$

$$\begin{aligned} C_3 = & 6\mu^2 - \frac{3k\mu^2}{\mu + \gamma_2} + \frac{3k p \mu^2}{\mu + \gamma_2} - \frac{3\alpha k p \mu^2}{\mu + \alpha\gamma_2} + 3\beta\eta_2\mu \\ & + 3\eta_2\mu + 3\alpha\gamma_2\mu + 3\gamma_2\mu - \frac{\beta k \eta_2 \mu}{\mu + \gamma_2} \\ & + \frac{\beta k \eta_2 p \mu}{\mu + \gamma_2} - \frac{\alpha k \gamma_2 \mu}{\mu + \gamma_2} + \frac{\alpha k p \gamma_2 \mu}{\mu + \gamma_2} \\ & - \frac{\alpha k \eta_2 p \mu}{\mu + \alpha\gamma_2} - \frac{\alpha k p \gamma_2 \mu}{\mu + \alpha\gamma_2} + \beta\eta_2^2 + \alpha\gamma_2^2 \\ & + \alpha\eta_2\gamma_2 + \alpha\beta\eta_2\gamma_2 + \beta\eta_2\gamma_2 + \eta_2\gamma_2 \\ C_4 = & 4\mu^3 - \frac{3k\mu^3}{\mu + \gamma_2} + \frac{3k p \mu^3}{\mu + \gamma_2} - \frac{3\alpha k p \mu^3}{\mu + \alpha\gamma_2} + 3\beta\eta_2\mu^2 \\ & + 3\eta_2\mu^2 + 3\alpha\gamma_2\mu^2 + 3\gamma_2\mu^2 \\ & - \frac{2\beta k \eta_2 \mu^2}{\mu + \gamma_2} + \frac{2\beta k \eta_2 p \mu^2}{\mu + \gamma_2} - \frac{2\alpha k \gamma_2 \mu^2}{\mu + \gamma_2} \\ & + \frac{2\alpha k p \gamma_2 \mu^2}{\mu + \gamma_2} - \frac{2\alpha k \eta_2 p \mu^2}{\mu + \alpha\gamma_2} - \frac{2\alpha k p \gamma_2 \mu^2}{\mu + \alpha\gamma_2} \\ & + 2\alpha\gamma_2^2\mu + 2\beta\eta_2^2\mu + 2\alpha\eta_2\gamma_2\mu \\ & + 2\alpha\beta\eta_2\gamma_2\mu + 2\beta\eta_2\gamma_2\mu + 2\eta_2\gamma_2\mu \\ & - \frac{\alpha\beta k \eta_2 \gamma_2 \mu}{\mu + \gamma_2} + \frac{\alpha\beta k \eta_2 p \gamma_2 \mu}{\mu + \gamma_2} \\ & - \frac{\alpha k \eta_2 p \gamma_2 \mu}{\mu + \alpha\gamma_2} + \alpha\eta_2\gamma_2^2 + \alpha\beta\eta_2\gamma_2^2 \\ & + \alpha\beta\eta_2^2\gamma_2 + \beta\eta_2^2\gamma_2 \end{aligned}$$

$$\begin{aligned} C_5 = & -\frac{k\mu^4}{\mu + \gamma_2} + \frac{k p \mu^4}{\mu + \gamma_2} - \frac{\alpha k p \mu^4}{\mu + \alpha\gamma_2} + \mu^4 + \beta\eta_2\mu^3 \\ & + \eta_2\mu^3 + \alpha\gamma_2\mu^3 + \gamma_2\mu^3 - \frac{\beta k \eta_2 \mu^3}{\mu + \gamma_2} \\ & + \frac{\beta k \eta_2 p \mu^3}{\mu + \gamma_2} - \frac{\alpha k \gamma_2 \mu^3}{\mu + \gamma_2} + \frac{\alpha k p \gamma_2 \mu^3}{\mu + \gamma_2} \\ & - \frac{\alpha k \eta_2 p \mu^3}{\mu + \alpha\gamma_2} - \frac{\alpha k p \gamma_2 \mu^3}{\mu + \alpha\gamma_2} + \beta\eta_2^2\mu^2 \\ & + \alpha\gamma_2^2\mu^2 + \alpha\eta_2\gamma_2\mu^2 + \alpha\beta\eta_2\gamma_2\mu^2 \\ & + \beta\eta_2\gamma_2\mu^2 + \eta_2\gamma_2\mu^2 - \frac{\alpha\beta k \eta_2 \gamma_2 \mu^2}{\mu + \gamma_2} \\ & + \frac{\alpha\beta k \eta_2 p \gamma_2 \mu^2}{\mu + \gamma_2} - \frac{\alpha k \eta_2 p \gamma_2 \mu^2}{\mu + \alpha\gamma_2} \\ & + \alpha\eta_2\gamma_2^2\mu + \alpha\beta\eta_2\gamma_2^2\mu + \alpha\beta\eta_2^2\gamma_2\mu \\ & + \beta\eta_2^2\gamma_2\mu + \alpha\beta\eta_2^2\gamma_2^2 \end{aligned}$$

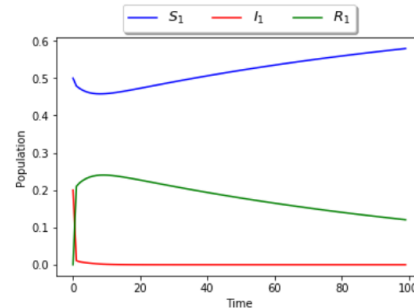
Dựa theo tiêu chuẩn ổn định Routh-Hurwitz, khi $R_0 > 1$, điểm cân bằng bệnh E_e ổn định tiệm cận.

V. KHẢO SÁT SỐ

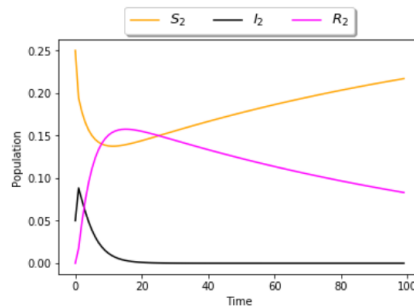
A. Sự lan truyền tắt dần

Trong kịch bản 1, chúng tôi chọn bộ tham số sau: $p = 0.7$, $\alpha = 0.2$, $k = 0.9$, $\beta = 3$, $\mu = 0.008$, $\eta_2 = 0.35$. Mật độ ban đầu của mỗi nhóm: $S_1 = 0.5$, $I_1 = 0.2$, $R_1 = 0$, $S_2 = 0.25$, $I_2 = 0.05$, $R_2 = 0$.

Trong ví dụ này ta có thể tính hệ số lây nhiễm cơ bản $R_0 \approx 0.87328$ dựa trên biểu thức (15). Hình 1 và 2 cho biết rằng thành phần nhiễm bệnh $I_1(t)$ và $I_2(t)$ tiến dần về 0 khi $t \rightarrow \infty$.



Hình 1: Diễn biến của S_1, I_1, R_1 theo thời gian t trong trường hợp $R_0 < 1$



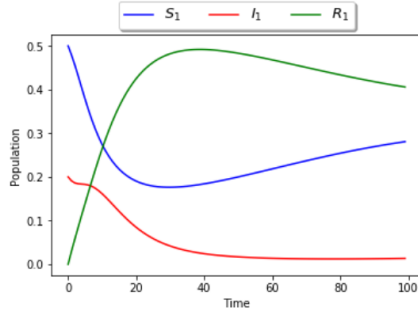
Hình 2: Diễn biến của S_2, I_2, R_2 theo thời gian t trong trường hợp $R_0 < 1$

Nhận xét: trong kịch bản 1, đồ thị thu được sau khi chạy mô hình đề xuất của chúng tôi xác định rằng mã độc sẽ biến mất sau một khoảng thời gian khi t tiến dần về dương vô cùng do tốc độ lây nhiễm lên các nút nhạy cảm không đuổi kịp tốc độ phục hồi của các nút đã nhiễm mã độc.

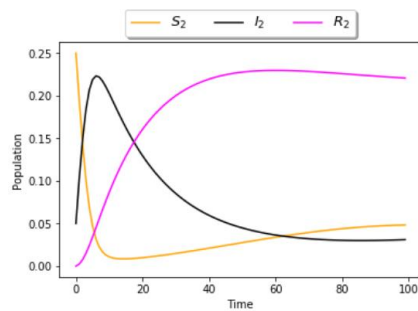
B. Sự lan truyền vẫn còn

Sau khi chạy mô hình với các tham số thỏa mãn R_0 phải nhỏ hơn một, chúng tôi tiếp tục thực hiện mô phỏng mô hình này với các tham số để giá trị hệ số lây nhiễm cơ bản R_0 phải lớn hơn một để kiểm tra tính đúng đắn của mô hình một cách toàn diện. Bộ tham số cho kịch bản 2: $p = 0.7$, $\alpha = 0.2$, $k = 0.9$, $\beta = 3$, $\mu = 0.008$, $\eta_2 = 0.05$. Mật độ ban đầu của mỗi nhóm: $S_1 = 0.5$, $I_1 = 0.2$, $R_1 = 0$, $S_2 = 0.25$, $I_2 = 0.05$, $R_2 = 0$.

Trong ví dụ này ta có thể tính hệ số lây nhiễm cơ bản $R_0 \approx 5.45264$ dựa trên biểu thức (15). Hình 3 và 4 cho biết rằng thành phần nhiễm bệnh $I_1(t)$ và $I_2(t)$ vẫn có giá trị dương khi $t \rightarrow \infty$.



Hình 3: Diễn biến của S_1, I_1, R_1 theo thời gian t trong trường hợp $R_0 > 1$



Hình 4: Diễn biến của S_2, I_2, R_2 theo thời gian t trong trường hợp $R_0 > 1$

Nhận xét: trong kịch bản 2, đồ thị thu được sau khi chạy mô hình đề xuất của chúng tôi xác định rằng mã độc sẽ vẫn tồn tại cho dù t tiến về vô cùng khi hệ số lây nhiễm cơ bản lớn hơn 1. Kết hợp với kết quả thu được trong kịch bản 1 đã chứng minh được tính đúng đắn về mặt toán học của mô hình mà chúng tôi đề xuất với của hệ số R_0 đã được chứng minh về mặt thực nghiệm là hoàn toàn chính xác.

VI. KẾT LUẬN

Bài viết này đã đề xuất thành công mô hình truyền nhiễm mới nhằm dự đoán sự lây lan của mã độc trong mạng cảm biến không dây dựa trên bản chất của chúng. Nó đã thành công thay thế những mô hình đơn truyền thống và chúng tôi cũng đã chạy thử nghiệm trên một vài kịch bản để kiểm tra độ ổn định của mô hình cũng như tìm được hệ số lây nhiễm cơ bản dựa trên phương pháp sử dụng ma trận thế hệ tiếp theo. Cuối cùng chúng tôi đã thực hiện mô phỏng số để quan sát sự lây lan của mã độc để xác minh kết quả lý thuyết thu được là đúng đắn. Do thiếu dữ liệu thực nghiệm nên bài viết này chỉ thực hiện mô phỏng số với các tham số được ước tính đơn giản, đây cũng là hạn chế của bài báo này. Trong những nghiên cứu tiếp theo, chúng tôi ngoài việc cải tiến mô hình này là thực hiện điều tra, thu

thập dữ liệu thực tế để làm các thực nghiệm nhằm mô tả chính xác hơn sự lây lan của mã độc.

TÀI LIỆU THAM KHẢO

- [1] Kermack, W.O., & McKendrick, A.G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, 115, 700-721.
- [2] Youssef, M., & Scoglio, C. (2011). An individual-based approach to SIR epidemics in contact networks. *Journal of theoretical biology*, 283(1), 136-144.
- [3] A. Martín del Rey, R. Casado Vara, S. Rodríguez González, A computational propagation model for malware based on the SIR classic model, *Neurocomputing*, Volume 484, 2022, Pages 161-171, ISSN 0925-2312.
- [4] Prajapati, A. (2023). A Propagation Model of Malicious Objects via Removable Devices and Sensitivity Analysis of the Parameters. In: Abd El-Latif, A.A., Maleh, Y., Mazurczyk, W., ELAffendi, M., I. Alkanhal, M. (eds) *Advances in Cybersecurity, Cybercrimes, and Smart Emerging Technologies. CCSET 2022. Engineering Cyber-Physical Systems and Critical Infrastructures*, vol 4. Springer, Cham.
- [5] S. Hosseini, A. Zandvakili, 2022, The SEIRS-C model of information diffusion based on rumour spreading with fuzzy logic in social networks, *International Journal of Computer Mathematics*, 99(9), pp. 1918-1940.
- [6] Liu, J., Paré, P. E., Nedić, A., Tang, C. Y., Beck, C. L., & Başar, T. (2016). On the analysis of a continuous-time bi-virus model. In 2016 IEEE 55th Conference on Decision and Control, CDC 2016 (pp. 290-295). Article 7798284 (2016 IEEE 55th Conference on Decision and Control, CDC 2016). Institute of Electrical and Electronics Engineers Inc.
- [7] Diekmann, O., Heesterbeek, J. A., & Roberts, M. G. (2010). The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society, Interface*, 7(47), 873-885.
- [8] van den Driessche, P., & Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180, 29-48.
- [9] Pastor-Satorras, R., & Vespignani, A. (2000). Epidemic spreading in scale-free networks. *Physical review letters*, 86 14, 3200-3.
- [10] Hai-Feng Huo, Peng Yang, Hong Xiang, Dynamics for an SIRS epidemic model with infection age and relapse on a scale-free network, *Journal of the Franklin Institute*, Volume 356, Issue 13, 2019, Pages 7411-7443, ISSN 0016-0032.
- [11] Li, C. H., & Yousef, A. M. (2019). Bifurcation analysis of a network-based SIR epidemic model with saturated treatment function. *Chaos (Woodbury, N.Y.)*, 29(3), 033129. <https://doi.org/10.1063/1.5079631>
- [12] Xinwei Wang, Haijun Peng, Boyang Shi, Dianheng Jiang, Sheng Zhang, Biaosong Chen, Optimal vaccination strategy of a constrained time-varying SEIR epidemic model, *Communications in Nonlinear Science and Numerical Simulation*, Volume 67, 2019, Pages 37-48, ISSN 1007-5704.
- [13] Yonghua Zheng, Jianhua Zhu, Chaoan Lai, "A SEIQR Model considering the Effects of Different Quarantined Rates on Worm Propagation in Mobile Internet", *Mathematical Problems in Engineering*, vol. 2020, Article ID 8161595, 16 pages, 2020.
- [14] Hernandez Guillen JD, Martin del Rey A, Casado-Vara R. Propagation of the Malware Used in APTs Based on Dynamic Bayesian Networks. *Mathematics*. 2021; 9(23):3097. <https://doi.org/10.3390/math9233097>

- [15] K. Khatwani, "On Routh-Hurwitz criterion," in IEEE Transactions on Automatic Control, vol. 26, no. 2, pp. 583-584, April 1981, doi: 10.1109/TAC.1981.1102670.
- [16] Gopal, M. (2002). Control Systems: Principles and Design, 2nd Ed. Tata McGraw-Hill Education. p. 14. ISBN 0070482896.
- [17] M. A. Choghadi and H. A. Talebi, "The Routh-Hurwitz Stability Criterion, Revisited: The Case of Multiple Poles on Imaginary Axis," in IEEE Transactions on Automatic Control, vol. 58, no. 7, pp. 1866-1869, July 2013, doi: 10.1109/TAC.2013.2242591.

PREDICTION MODEL FOR MALWARE IN WIRELESS SENSOR NETWORK

Abstract: The propagation model is a mathematical model that has been and is currently under investigation in various fields, most notably epidemiology, statistics, etc. Based on the similarities between biological virus and malware, this article introduces a model were described by means of a ODEs. It predicts malware propagation based on the nature of malware. Its objects may include enterprises or personal computers. In comparison to a simple epidemic model, we proposed a new model in which the nodes are closely related through the transmission of malware between them. This allows us to promptly implement policies to effectively control the spread of malware.

Keywords: wireless sensor network, malware, characteristic equation, system of differential equations.



Trần Hải Anh (1993). Tốt nghiệp đại học ngành Công nghệ thông tin tại Học viện Quản lý Giáo dục. Hiện đang là học viên cao học ngành Khoa học máy tính tại Học viện Công nghệ Bưu Chính Viễn Thông. Lĩnh vực nghiên cứu: giải tích số, mô hình hóa và mô phỏng, cấu trúc đồ thị, học máy thống kê, học sâu và hệ tư vấn.



Đỗ Xuân Chợt (1985). Nhận bằng thạc sĩ ngành Khoa học máy tính và cơ sở máy tính năm 2010, nhận bằng Tiến sĩ ngành Tự động hóa thiết kế tại Đại học Kỹ thuật điện quốc gia Saint Petersburg. Năm 2022, ông được bổ nhiệm làm phó giáo sư chuyên ngành an toàn thông tin và hiện đang là giảng viên Khoa An toàn thông tin Học viện Công nghệ Bưu chính Viễn thông và FPTU tại Việt Nam. Lĩnh vực quan tâm: bảo mật thông tin, mô hình hóa, hệ thống điều khiển, thuật toán hóa.



Nguyễn Đức Việt (1969) Nhận học vị Tiến sĩ năm 2016 Tại Học Viện Kỹ Thuật Quân Sự, Hiện công tác tại Học viện Công nghệ Bưu chính Viễn thông. Lĩnh vực nghiên cứu: Định vị vô tuyến, hệ thống giám sát Multilateration, Hệ thống dẫn đường và giám sát phương tiện bay, IoT và các ứng dụng