

# MỤC LỤC

Danh sách hình vẽ

Danh sách bảng

Mở đầu	1
<b>Chương 1. GIỚI THIỆU VỀ MẠNG XÃ HỘI VÀ TÁC HẠI CỦA THÔNG TIN SAI LỆCH</b>	<b>9</b>
1. Giới thiệu chung về mạng xã hội	9
1.1. Đặc điểm của MXH	9
1.2. Lợi ích và các tác động của MXH	10
2. Sự lây lan và tác hại của thông tin sai lệch trên MXH	12
2.1. Định nghĩa thông tin sai lệch và tính chất	13
2.2. Thực trạng sự phát tán của thông tin sai lệch trên mạng xã hội	14
Kết luận chương	17
<b>Chương 2. THÔNG TIN SAI LỆCH VÀ CƠ CHẾ LAN TRUYỀN THÔNG TIN SAI LỆCH</b>	<b>18</b>
1. Mô hình lan truyền thông tin	18
2. Cơ chế lan truyền thông tin sai lệch trên hai mô hình thông tin	20
2.1. Mô hình tầng độc lập	20
2.2. Mô hình ngưỡng tuyến tính	21
<b>Chương 3. GIẢI PHÁP NGĂN CHẶN SỰ PHÁT TÁN CỦA THÔNG TIN SAI LỆCH</b>	<b>24</b>
1. Bài toán 1: Ngăn chặn thông tin sai lệch với ngân sách giới hạn	25
1.1. Tổng quan bài toán	25
1.2. Mô hình lan truyền thông tin Ngưỡng tuyến tính với bước thời gian rời rạc (Time Constraint Deterministic Linear Threshold - T-DLT).	26
1.3. Định nghĩa bài toán	28
1.4. Độ phức tạp của bài toán LSE	29
1.5. Thuật toán đề xuất	30
1.6. Kết quả thực nghiệm với dữ liệu đã công bố	33
2. Bài toán 2: Tìm chi phí nhỏ nhất để ngăn chặn thông tin sai lệch	38
2.1. Tổng quan bài toán	38
2.2. Mô hình lan truyền thông tin Ngưỡng tuyến tính Linear Threshold	39
2.3. Định nghĩa bài toán	39
2.4. Độ phức tạp của bài toán	41
2.5. Thuật toán đề xuất	43

2.6. Kết quả thực nghiệm với dữ liệu đã công bố . . . . .	48
<b>Chương 4. ÁP DỤNG CÁC GIẢI PHÁP ĐỀ XUẤT TRÊN DỮ LIỆU THỰC</b>	<b>51</b>
1. Thu thập dữ liệu . . . . .	52
1.1. Xác định nguồn phát tán thông tin sai lệch . . . . .	52
1.2. Xây dựng công cụ thu thập dữ liệu trên mạng xã hội Facebook . . .	53
2. Mô hình hóa dữ liệu thu được . . . . .	56
3. Áp dụng kịch bản và giải pháp ngăn chặn . . . . .	59
4. Ảnh xạ người dùng thực . . . . .	64
5. Giải pháp ngăn chặn . . . . .	65
<b>Tài liệu tham khảo</b>	<b>69</b>

## DANH SÁCH HÌNH VẼ

2.1	Một số ví dụ quá trình lan truyền thông tin trên mô hình IC . . . . .	21
2.2	Ví dụ quá trình lan truyền thông tin trên mô hình LT . . . . .	22
3.1	Phép dẫn đa thức từ bài toán SC về bài toán LSE. . . . .	30
3.2	So sánh chất lượng lời giải và thời gian chạy của các thuật toán khi $k$ thay đổi, $d = 5$ , $\theta = 0.5$ . . . . .	36
3.3	So sánh chất lượng lời giải và thời gian chạy của các thuật toán khi $\theta$ thay đổi, $d = 5$ , $k = 50$ . . . . .	37
3.4	So sánh chất lượng lời giải và thời gian chạy của các thuật toán khi $d$ thay đổi, $\theta = 0.5$ , $k = 50$ trong mạng Gnutella. . . . .	38
3.5	Giản thể từ bài toán $s$ - $t$ paths sang bài toán TMB. . . . .	41
3.6	Ví dụ tạo ra một cây gốc $I$ từ $G$ trên mô hình LT . . . . .	45
3.7	So sánh chất lượng lời giải và thời gian chạy của các thuật toán cho bài toán TMB . . . . .	50
3.8	Kiểm tra kết quả của thuật toán STMB . . . . .	50
4.1	Tổng thể giải pháp . . . . .	51
4.2	Mô phỏng cấu trúc toàn mạng dữ liệu thu được . . . . .	57
4.3	So sánh chất lượng lời giải của các thuật toán với dữ liệu Facebook . . . . .	60
4.4	Kết quả mô phỏng với dữ liệu thu thập được trên Facebook của thuật toán FLE . . . . .	61
4.5	So sánh chất lượng lời giải của các thuật toán với dữ liệu thu thập được cho bài toán TMB . . . . .	62
4.6	So sánh thời gian thực hiện của các thuật toán với dữ liệu thu thập được cho bài toán TMB . . . . .	63
4.7	Kết quả mô phỏng với dữ liệu thu thập được trong trường hợp $\gamma = 15$ của thuật toán STMB . . . . .	64

## DANH SÁCH BẢNG

3.1	Các bộ dữ liệu . . . . .	34
3.2	Các bộ dữ liệu dùng trong thực nghiệm với giải pháp TMB . . . . .	48
3.3	So sánh thời gian chạy của các thuật toán với trường hợp $\gamma$ lớn nhất . . . . .	50
4.1	Ảnh xạ ID người dùng sang chỉ số được đánh số - ID.txt . . . . .	56
4.2	Mô tả cụ thể của đồ thị MXH thu được - Network.txt . . . . .	57
4.3	Dữ liệu đầu vào ứng với giải pháp FLE - input.txt . . . . .	58
4.4	Dữ liệu đầu vào ứng với giải pháp STMB - TestFacebook.txt . . . . .	59
4.5	Kết quả thực hiện giải pháp FLE với dữ liệu thu thập - Output.txt . . . . .	59
4.6	Kết quả thực hiện giải pháp SMTB với dữ liệu thu thập - Result.txt . . . . .	62
4.7	Danh sách người dùng thực - Kết quả giải pháp FLE . . . . .	65
4.8	Danh sách người dùng thực - Kết quả giải pháp SMTB . . . . .	65

## DANH MỤC CÁC TỪ VIẾT TẮT

Từ viết tắt	Ý nghĩa
APT	Tấn công mạng sử dụng công nghệ cao - Advanced Persistent Threat
FLE	Thuật toán nhanh và hiệu quả để giới hạn sự lây nhiễm thông tin - Fast And Effective Limiting Epidemics
IC	Mô hình tầng độc lập - Independent Cascade
IP	Giao thức Internet - Internet Protocol
LSE	Hạn chế sự lây lan của thông tin sai lệch - Limiting the Spread of Epidemics
LT	Mô hình ngưỡng tuyến tính - Linear Threshold
MXH	Mạng xã hội
T-DLT	Mô hình ngưỡng tuyến tính với bước thời gian rời rạc - Time-Discrete Linear Threshold
TMB	Xác định và ngăn chặn thông tin sai lệch trên mạng xã hội - Targeted Misinformation Blocking

# MỞ ĐẦU

## 1. Lí do chọn đề tài

Cùng với sự phát triển của Internet, các mạng xã hội đã phát triển mạnh mẽ và trở thành một xu hướng mới thu hút nhiều người sử dụng trên toàn thế giới. Theo những khảo sát gần đây, có gần một nửa dân số thế giới, tức là hơn 3 tỷ người sử dụng mạng xã hội. Nhờ có mạng xã hội, người dùng có thể trao đổi thông tin với nhau một cách nhanh chóng bất kể khoảng cách về địa lý và thời gian. Bên cạnh đó, mạng xã hội còn cung cấp cho người dùng rất nhiều ứng dụng hữu ích, làm cho cuộc sống của con người ngày càng trở nên thuận tiện hơn. Ngoài những đặc tính kế thừa của mạng lưới xã hội thực như: tương tác giữa người dùng, lan truyền thông tin, tạo ảnh hưởng trong cộng đồng thì mạng xã hội còn mang nhiều đặc tính mới như: cập nhật thông tin thực lên mạng xã hội một cách nhanh chóng, sự lan truyền thông tin giữa người dùng xảy ra trong thời gian ngắn, sự bùng nổ thông tin với các nguồn tin tức khác nhau, v.v... Có thể nói, hiện nay mạng xã hội đang từng bước trở thành một kho tri thức mới mà con người có thể dễ dàng tiếp cận.

Tuy vậy, đi cùng với những lợi ích mạng xã hội mang lại thì còn nhiều nguy cơ và hiểm họa tới người dùng. Một trong các yếu tố mang lại nhiều tác hại cho cộng đồng người dùng là sự phát tán thông tin sai lệch. Các yếu tố này gây ra những tác hại lớn đối với cộng đồng người sử dụng mạng xã hội. Không chỉ ở Việt Nam, những tác hại này diễn ra trên phạm vi toàn thế giới. Việc lan truyền thông tin sai lệch có thể gây ra những tác động xã hội tiêu cực, thậm chí là tổn thất lớn về kinh tế. Chẳng hạn, những tin đồn không hay về sức khỏe của tổng thống Mỹ ông Obama vào tháng 4 năm 2013 đã dẫn đến sự bất ổn của thị trường tài chính ở phố Wall. Trước thềm Đại hội đại biểu toàn quốc lần thứ XII của Đảng Cộng sản Việt Nam, đã có những thông tin về gia đình cũng như tài sản của Thủ tướng Nguyễn Xuân Phúc lúc đó vẫn còn là Phó Thủ tướng Chính phủ, nhằm hạ uy tín gây mất lòng tin của nhân dân đối với Thủ tướng. Trước những thực trạng đó, ngăn chặn, phòng chống lây lan thông tin sai lệch trở thành mục tiêu hàng đầu của các quốc gia.

Các mạng xã hội lớn như Twitter, Youtube và Facebook đang phải đối mặt với việc gia tăng áp lực từ chính phủ của nhiều nước đe dọa áp đặt các đạo luật mới và khoản phạt nếu các mạng xã hội này không có biện pháp loại bỏ nhanh chóng các nội dung tuyên truyền cực đoan cũng như các nội dung vi phạm luật pháp các nước sở tại.

Ở Việt Nam, trong thời gian gần đây, Đảng và Nhà nước ta đã đề ra rất nhiều biện pháp nhằm ngăn chặn thông tin sai lệch trên mạng xã hội và đặc biệt là bộ Thông tin và truyền thông đã tiến hành ngăn chặn, xử lý những thông tin sai sự thật, nghiêm túc xem

xét đánh giá, xử lý sai phạm, bộ Thông tin và truyền thông đã bắt các nhà MXH gỡ bỏ các thông tin sai lệch được đăng tải [1]. Bộ quan tâm trước hết đến báo chí và chất lượng báo chí, có những bài viết với số lượng nhiều nhưng chất lượng kém. Bên cạnh đó, ta chủ động phát hiện những thông tin sai lệch, kịp thời chỉnh sửa, đính chính, đưa ra thông tin chính xác nhất.

Ngoài ra, những thông tin sai lệch, không kiểm chứng, giật tít tạo sự tò mò cho người dùng trên mạng xã hội còn tiềm ẩn các nguy cơ phát tán mã độc, lừa đảo trên mạng gây mất an ninh, an toàn thông tin.

Với thực trạng đó, các nhà khoa học đã nghiên cứu những giải pháp hiệu quả để ngăn chặn thông tin sai lệch. Trong đó việc mô hình hóa quá trình lan truyền thông tin trên mạng là nền tảng trong tiếp cận của họ. Các mô hình lan truyền thông tin, hay còn gọi là mô hình khuếch tán thông tin được các nhà khoa học đưa ra đã mô hình hóa toán học và mô tả một cách tương đối chính xác một mạng lưới và quá trình lan truyền thông tin, lan truyền dịch bệnh ở trên mạng lưới đó.

Tuy nhiên, những nghiên cứu đó còn gặp phải một số hạn chế như chỉ thực hiện trên các dữ liệu có sẵn, chỉ tập trung vào nghiên cứu lý thuyết, và đặc biệt chưa có nghiên cứu nào thực hiện được đối với bài toán ngăn chặn thông tin sai lệch tại Việt Nam.

Thúc đẩy bởi những hiện tượng trên và yêu cầu bức thiết của việc giải quyết và ngăn chặn những tác hại do tin đồn trên mạng xã hội mang lại. Nhóm tác giả đã mạnh dạn nghiên cứu đề tài nhằm mục đích tìm ra giải pháp hợp lý để giải quyết những vấn đề nêu trên. Đặc biệt là trong bối cảnh ở Việt Nam ngày càng mở rộng hợp tác quan hệ với các nước trên thế giới, đi kèm theo nhiều lợi ích đó là đối mặt với nhiều nguy cơ thách thức về An ninh quốc gia. Đặc biệt là các thế lực thù địch lợi dụng MXH để thực hiện những âm mưu, hoạt động nhằm xuyên tạc, chống phá đường lối, chính sách của Đảng, pháp luật của Nhà nước. Nghiêm trọng hơn, trong thời gian gần đây, chúng thực hiện những thủ đoạn cực kỳ tinh vi: thuê người làm dữ liệu giả (ảnh, video clip, audio hiện trường) để đưa lên mạng xã hội, các diễn đàn, xuyên tạc sự thật, cung cấp thông tin sai lệch về các cá nhân, các cán bộ, Đảng viên; dùng kỹ thuật và công nghệ để chỉnh sửa dữ liệu cũ, chỉnh sửa hình ảnh, cắt ghép, tạo bằng chứng và thông tin giả; tự bịa ra các bài phỏng vấn nhân vật, sự kiện, bịa đặt các trang hồ sơ liên quan đến các nhân vật nổi tiếng, các nhà lãnh đạo, thân nhân của họ và kích thích trí tò mò của công chúng bằng “thông tin lẻ trái, thông tin bí mật”, những việc làm này gây hậu quả vô cùng nghiêm trọng đối với tình hình An ninh quốc gia của Việt Nam.

## **2. Tình hình nghiên cứu liên quan đến đề tài**

Ngăn chặn sự phát tán của thông tin sai lệch, tin đồn hiện nay là một chủ đề nóng

được nhiều nhà khoa học quan tâm. Rất nhiều những nghiên cứu và bằng sáng chế liên quan vấn đề này có tính ứng dụng cao được áp dụng trên các mạng xã hội, ứng dụng trong giáo dục, y tế sức khỏe, ứng dụng cho chính phủ cũng như các vấn đề bảo vệ An ninh Quốc gia.

Không chỉ ở Việt Nam, đối với các nước trên thế giới, sự phát tán của thông tin sai lệch hiện nay đã trở thành vấn nạn nghiêm trọng. Một số nước đã thành lập các trung tâm chống tin giả sử dụng các biện pháp, giải pháp xã hội cũng như kỹ thuật để ngăn chặn, hạn chế sự phát tán của thông tin sai lệch tới công chúng.

Nguồn phát tán thông tin sai lệch có thể được phát hiện thông qua khảo sát người dùng hoặc các phương pháp khai phá dữ liệu được phát hiện thông qua việc sử dụng các đặc trưng như: thời gian, cấu trúc và ngữ nghĩa [2]. Về phương diện mô hình hóa toán học, để phát hiện được nguồn phát tán thông tin sai lệch, Zhang và cộng sự đề xuất vấn đề Time Constrained Misinformation Detection (Phát hiện thông tin sai lệch trong thời gian ràng buộc - TCMD) tìm kiếm tập đỉnh nhỏ nhất sao cho phát hiện được thông tin sai lệch lớn hơn ngưỡng  $t$ . Christakis [3] và cộng sự đề xuất một phương thức phát hiện sự lan truyền của “ô dịch” bằng cách theo dõi người dùng sử dụng phương pháp lựa chọn ngẫu nhiên.

Trong việc đấu tranh với sự phát tán của thông tin sai lệch, nhiều tác giả đã đề xuất chiến lược “khử nhiễm”. Điển hình là nghiên cứu của Nguyen và cộng sự [2] đề xuất chiến lược cho việc “khử nhiễm” thông tin sai lệch trên mạng xã hội bằng cách dùng thông tin tốt, chính thống để “khử nhiễm” thông tin xấu. Tuy nhiên, khi người dùng đã tin theo thông tin sai lệch sẽ là khó để thay đổi quan điểm của họ và chúng ta cần biết nội dung của thông tin xấu để đưa ra những thông tin chính thống phù hợp.

Gần đây, các nhà khoa học đã đề xuất một chiến lược chung trong việc ngăn chặn sự phát tán của tin đồn, thông tin sai lệch, đó là ngăn chặn những tài khoản và liên kết có vai trò quan trọng trong quá trình lan truyền, có một số phương pháp heuristic hiệu quả đã được đề xuất. Khalil [4] và cộng sự đề xuất bài toán xóa bỏ tập cạnh để giảm thiểu ảnh hưởng của tập nguồn, họ cũng đề xuất một thuật toán giải quyết có tỷ lệ xấp xỉ  $1 - 1/e - \varepsilon$ . Ceren Budak [5] nghiên cứu bài toán hạn chế sự lây lan của một chiến dịch lan truyền thông tin “xấu” trong mạng xã hội, dựa trên các mô hình lan truyền và tính cạnh tranh của thông tin trên mạng xã hội trực tuyến.

Canh.V Pham [6] và các cộng sự nghiên cứu việc chặn thông tin sai lệch trong một thời gian và ngân sách cho trước. Cui [7] và cộng sự tập trung vào việc nghiên cứu lựa chọn những điểm đặt các cảm biến sao cho phát hiện sự bùng nổ thông tin với xác suất cao nhất. Zhang đề xuất vấn đề  $t$  - Monitors Placement ngăn ngừa việc lan truyền thông tin sai lệch với ngưỡng đảm bảo  $t \in [0,1]$  đề xuất chiến lược tiêm vaccin cho  $k$  đỉnh còn



lại sao cho số đỉnh bị lây lan là nhỏ nhất trong mô hình IC. Các tác giả cho thấy rằng vấn đề tối ưu hóa này là NP-khó ngay cả việc tính toán hàm mục tiêu. Ngoài ra, có rất nhiều nghiên cứu liên quan đến vấn đề này trong thời gian gần đây.

Tuy nhiên, hầu hết các nghiên cứu trên đều được thực hiện trên dữ liệu có sẵn đã được công bố và chúng bị hạn chế theo những khía cạnh sau:

- Việc tính toán hàm mục tiêu là NP-Khó trên các mô hình phát tán thông tin cổ điển dẫn đến việc khó thực hiện giải pháp trên các mạng xã hội thực.
- Yếu tố thời gian chưa được tập trung xem xét trong ngăn chặn thông tin sai lệch. Đây là điều rất quan trọng vì thông tin chỉ phát tán nhanh trong một khoảng thời gian nhất định, nếu thông tin sai lệch phát tán trong thời gian càng dài thì càng gây nhiều thiệt hại và ảnh hưởng tới công chúng. Hơn nữa, việc ngăn chặn thông tin sai lệch với nguồn lực cho trước có một sự hạn chế là ta không biết ngân sách bao nhiêu để ngăn chặn sự bùng phát của thông tin sai lệch.
- Các nghiên cứu mới chỉ dừng lại ở mô hình và thuật toán mà chưa kết hợp với việc thu thập dữ liệu hiện thực, đặc biệt chưa áp dụng ở các mạng xã hội ở Việt Nam.

Trong đề tài này, nhóm tác giả tập trung khắc phục các hạn chế trên của những nghiên cứu đã có. Đặc biệt, nhóm tác giả đã thu thập dữ liệu của các đối tượng phát tán thông tin sai lệch ở Việt Nam và đưa ra giải pháp ngăn chặn thông tin sai lệch trên các dữ liệu được thu thập đó. Qua đó có thể xây dựng một hệ thống ngăn chặn thông tin sai lệch trên dữ liệu thu thập được bằng việc đưa ra danh sách người dùng mà việc ngăn chặn họ tin theo, nghe theo, phát tán thông tin sai lệch sẽ hạn chế được tối đa thông tin sai lệch.

### 3. Mục tiêu, nhiệm vụ nghiên cứu

- **Mục tiêu nghiên cứu:** Đề tài thực hiện để đạt được các mục tiêu sau:
  - + Tìm hiểu, làm rõ tác hại của thông tin sai lệch trên các mạng xã hội.
  - + Tìm hiểu cơ chế lây lan tin đồn trên mạng xã hội. Qua đó đề xuất mô hình lan truyền thông tin phù hợp có yếu tố thời gian.
  - + Thiết lập bài toán hạn chế tin đồn trên mạng xã hội trên mô hình lan truyền thông tin mới.
  - + Đưa ra giải pháp, thuật toán hiệu quả để ngăn chặn sự lây lan tin đồn trên mạng xã hội.
  - + Thu thập dữ liệu về người dùng trên mạng xã hội thực, đặc biệt là dữ liệu các đối tượng phát tán thông tin. Đưa ra kết quả thực nghiệm và đánh giá.

- **Nhiệm vụ nghiên cứu:** Để đạt được mục tiêu nghiên cứu, nhóm tác giả giải quyết lần lượt các nhiệm vụ sau:

- + Nghiên cứu mạng xã hội, cách thức lan truyền thông tin trên mạng xã hội, mô hình hóa toán học của mạng xã hội trong các nghiên cứu từ trước tới nay trên các hội nghị khoa học, tạp chí khoa học uy tín qua các tài liệu, bài báo cáo.
- + Nghiên cứu các thuật toán liên quan đã có giải quyết bài toán ngăn chặn sự lan truyền thông tin sai lệch trên mạng xã hội trực tuyến, cập nhật các nghiên cứu trong thời gian gần đây.
- + Đưa ra thuật toán hiệu quả cho bài toán ngăn chặn sự lây lan của tin đồn trên mạng xã hội, so sánh với các thuật toán mới nhất. Nghiên cứu mô hình hóa mạng xã hội và quá trình lan truyền thông tin trên mạng xã hội.
- + Nghiên cứu giải pháp lấy dữ liệu trên các mạng xã hội, trong phạm vi của đề tài là mạng xã hội Facebook.
- + Đề xuất giải pháp hạn chế thông tin sai lệch trên mạng xã hội và thực nghiệm kiểm tra giải pháp.
- + Thực nghiệm giải pháp trên dữ liệu thực thu thập từ mạng xã hội Facebook. Đưa ra danh sách người dùng có vai trò nhất trong việc phát tán thông tin sai lệch từ một người cho trước. Từ đó đưa ra những giải pháp thích hợp với những người dùng này.

#### **4. Đối tượng và phạm vi nghiên cứu của đề tài**

- **Đối tượng nghiên cứu:**

- + Các mạng xã hội trên thế giới đặc biệt là các mạng xã hội phổ biến ở Việt Nam được đa số người dùng sử dụng.
- + Hành vi phát tán thông tin trên mạng xã hội.
- + Các đối tượng phát tán thông tin sai lệch.

- **Phạm vi nghiên cứu**

- + Thời gian: Các mạng xã hội từ năm 2001 tới nay.
- + Không gian: Mạng Internet.

## 5. Phương pháp nghiên cứu

Trên cơ sở những điều kiện và đặc điểm nêu trên, để đạt được mục tiêu nghiên cứu của đề tài, nhóm tác giả sử dụng tổng hợp các cách tiếp cận và phương pháp nghiên cứu sau:

### - Phương pháp nghiên cứu tài liệu:

- + Phương pháp phân tích và tổng hợp lý thuyết: Nhóm tác giả đã tham khảo các bài báo, tạp chí, hội nghị có uy tín trong nước và thế giới về những nghiên cứu có liên quan đến đề tài, từ đó phân tích đánh giá, có cái nhìn gần hơn, ví dụ như :
  - Science direct: <http://www.sciencedirect.com/>.
  - IEEE Xplore Digital Library: <http://ieeexplore.ieee.org>.
  - The ACM Digital Library: <http://dl.acm.org/>.
  - Google Scholar: <https://scholar.google.com>.
- + Phương pháp phân loại và hệ thống hóa lý thuyết: Từ những lý thuyết, tài liệu đã có, nhóm tác giả phân loại, hệ thống hóa, sắp xếp một cách khoa học, dễ hiểu. Đồng thời, có những nhận định, đánh giá về những lý thuyết, công trình nghiên cứu trước đây, qua đó, nhóm tác giả có những đề xuất tìm ra phương pháp giải quyết tối ưu hơn cho đề tài.

### - Phương pháp nghiên cứu thực tiễn

- + Phương pháp hỏi chuyên gia: Thu thập thông tin tài liệu, ý kiến của chuyên gia, đánh giá vấn đề, hỏi ý kiến từ các chuyên gia hàng đầu về lĩnh vực công nghệ thông tin và an toàn thông tin để có cái nhìn tổng quát hơn về vấn đề đang nghiên cứu.
- + Từ thực tiễn đề xuất mô hình phù hợp với thực tế hơn.
- + Phương pháp mô hình hóa: Từ những mô hình mạng xã hội, nhóm tác giả mô hình hóa thành các bài toán phù hợp, từ đó sẽ giải quyết những yêu cầu của bài toán bằng các công cụ của toán học, khoa học máy tính,...
- + Phương pháp thống kê: Tổng hợp, thống kê những số liệu thu được, mô hình hóa bằng các biểu đồ để có cái nhìn tổng quát hơn, từ đó dễ dàng đánh giá, nhận định vấn đề, đánh giá thuật toán và mô hình đề xuất.

## 6. Ý nghĩa khoa học, ý nghĩa thực tiễn của đề tài

### - Ý nghĩa khoa học:

- + Cung cấp cái nhìn tổng quan về thông tin sai lệch và thực trạng lan truyền thông tin sai lệch hiện nay.
- + Đưa ra những lý thuyết mới có ý nghĩa, những giải pháp có tính khả thi trong nghiên cứu cơ bản cũng như thực tiễn hạn chế thông tin sai lệch trên mạng xã hội.
- + Đưa ra được mô hình mới và thuật toán mới khắc phục được những hạn chế của các mô hình và thuật toán đã công bố.

Qua đó cho thấy, một phần công trình nghiên cứu của nhóm được các đồng nghiệp trên thế giới đánh giá cao và có giá trị về mặt học thuật cũng như giá trị trong thực tiễn. Nhóm tác giả đã vận dụng những nghiên cứu quan trọng này trong đề tài, đề đưa ra giải pháp ngăn chặn thông tin sai lệch có hiệu quả.

#### - Ý nghĩa thực tiễn:

- + Giúp người dùng có cái nhìn rõ hơn về những tác hại của mạng Internet và đặc biệt là thực trạng về thông tin sai lệch.
- + Đánh giá thực trạng lan truyền, quy luật lan truyền của thông tin sai lệch trên mạng xã hội. Đề xuất một giải pháp hạn chế thông tin sai lệch trên mạng xã hội.
- + Là nền tảng để xây dựng một hệ thống để chặn thông tin sai lệch và có ý nghĩa trong thực tiễn. Đặc biệt trong bối cảnh hiện nay, sự phát tán rộng rãi của thông tin sai lệch là vấn nạn đối với các quốc gia. Một số nước phát triển đã xây dựng trung tâm phòng chống tin giả.
- + Từ việc ngăn chặn thông tin sai lệch, ta hạn chế được rất nhiều những hậu quả nghiêm trọng do thông tin sai lệch gây ra.

## 7. Cấu trúc của đề tài

Ngoài phần mở đầu, kết luận, danh mục tài liệu tham khảo và phụ lục, đề tài được cấu trúc gồm 4 chương:

### **Chương 1: Giới thiệu về mạng xã hội và tác hại của thông tin sai lệch**

Chương này giới thiệu tổng quan về mạng xã hội bao gồm: Định nghĩa mạng xã hội, những đặc trưng cơ bản, lợi ích và tác hại của mạng xã hội. Đặc biệt, trong chương này trình bày thực trạng lan truyền thông tin sai lệch trên mạng xã hội đối với thế giới và Việt Nam, từ đó đặt ra vấn đề cấp thiết trong việc tìm hiểu thông tin sai lệch, cơ chế lan truyền và giải pháp hạn chế tác hại của nó.

## **Chương 2: Thông tin sai lệch và cơ chế lan truyền thông tin sai lệch**

Chương này nhóm tác giả trình bày định nghĩa thông tin sai lệch, những nguy cơ và hậu quả do thông tin sai lệch gây ra đối với các cá nhân, tổ chức. Đồng thời, phân tích cơ chế lan truyền thông tin và những đặc tính của hai mô hình lan truyền thông tin đang được sử dụng rộng rãi bao gồm: Mô hình tầng độc lập và mô hình ngưỡng tuyến tính.

## **Chương 3: Giải pháp ngăn chặn sự phát tán của thông tin sai lệch**

Từ thực trạng đã nêu trong Chương 2 và xuất phát từ những công trình nghiên cứu liên quan trước đó, nhóm tác giả đã phát biểu tổng quát bài toán cực tiểu hóa thiệt hại do thông tin sai lệch gây ra trên mạng xã hội trực tuyến. Đề xuất hai kịch bản ngăn chặn có tính thực tiễn, tương ứng với hai bài toán tối ưu, nghiên cứu, chứng minh hai bài toán thuộc lớp #P – khó, NP - khó, đồng thời đề xuất các giải pháp hiệu quả để giải quyết hai bài toán này.

## **Chương 4: Áp dụng các giải pháp đề xuất trên dữ liệu thực**

Chương này nhóm tác giả trình bày một mô hình tổng quan gồm 4 bước nhằm giải quyết bài toán ngăn chặn thông tin sai lệch trên mạng xã hội. Đầu tiên là bước thu thập dữ liệu thực từ mạng xã hội Facebook tại Việt Nam với phạm vi xung quanh các đối tượng chủ mưu, cầm đầu các tổ chức phản động, các đối tượng có uy tín lớn. Sau đó nhóm tác giả tiến hành mô hình hóa toán học các dữ liệu thu được tương ứng với hai kịch bản, bài toán. Sau đó áp dụng hai giải pháp đề xuất, đưa ra tập đỉnh có vai trò quan trọng trong lan truyền thông tin trên đồ thị. Và bước cuối cùng, ánh xạ ngược để đưa ra tập các người dùng, tài khoản có ảnh hưởng, vai trò lớn đối với quá trình lan truyền thông tin sai lệch. Ngoài ra nhóm tác giả còn đề xuất các giải pháp để ngăn chặn, vô hiệu hóa hoạt động lan truyền thông tin này của các đối tượng, nhất là trong bối cảnh dự thảo Luật An ninh mạng đã được thông qua và Luật An ninh mạng đang được chỉnh lý, sắp sửa ban hành.

# **CHƯƠNG 1**

## **GIỚI THIỆU VỀ MẠNG XÃ HỘI VÀ TÁC HẠI CỦA THÔNG TIN SAI LỆCH**

### **1. Giới thiệu chung về mạng xã hội**

Mạng xã hội (MXH), hay còn gọi là mạng xã hội ảo (Social Network) là dịch vụ nối kết các thành viên cùng sở thích trên Internet với nhiều mục đích khác nhau không phân biệt không gian và thời gian. Những người tham gia vào MXH còn được gọi là “cư dân mạng”.

MXH bao gồm rất nhiều các dịch vụ mang các tính năng riêng biệt: chat, mail, blog, video, chia sẻ thông tin giúp kết nối người dùng dựa vào các mối quan hệ đặc trưng như: quan hệ họ hàng, sở thích, ý tưởng. Chính vì thế, mạng xã hội dễ dàng giúp con người có thể tìm kiếm kết nối với các mối quan hệ dựa trên các nhóm, trường, cơ quan, dựa trên các thông tin cá nhân, địa chỉ.

MXH được cấu thành dựa vào hai thành phần cơ bản. Thứ nhất là có sự tham gia trực tuyến của các cá nhân hay các chủ thể. Thứ hai là MXH sẽ có các trang web mở, người dùng tự xây dựng nội dung trong đó và các thành viên trong nhóm đây sẽ biết được các thông tin mà người dùng viết.

Hiện nay thế giới có rất nhiều mạng xã hội khác nhau, thị trường Bắc Mỹ và Tây Âu nổi tiếng với MySpace và Facebook; Nam Mỹ với Orkut và Hi5; Friendster tại Châu Á và các đảo quốc Thái Bình Dương. Các trang MXH khác gặt hái được thành công đáng kể theo vùng miền như Bebo tại Anh Quốc, CyWorld tại Hàn Quốc, Mixi tại Nhật Bản. Tại Việt Nam, theo một nghiên cứu mới đây của DoubleClick Ad Planner, các trang MXH có lượng truy cập gần 16 triệu lượt/tháng trong đó có ba MXH lớn nhất tại Việt Nam: Facebook, Zing Me và Yume.

Mục tiêu là tạo ra một hệ thống trên nền Internet cho phép người dùng giao lưu và chia sẻ thông tin một cách có hiệu quả, vượt ra ngoài những giới hạn về địa lý và thời gian. Xây dựng lên một mẫu định danh trực tuyến nhằm phục vụ những yêu cầu công cộng chung và những giá trị của cộng đồng. Qua đó, nhằm nâng cao vai trò của mỗi công dân trong việc tạo lập quan hệ và tự tổ chức xoay quanh những mối quan tâm chung trong những cộng đồng thúc đẩy sự liên kết các tổ chức xã hội.

#### **1.1. Đặc điểm của MXH**

MXH trên Internet bao gồm các đặc điểm nổi bật: Tính liên kết cộng đồng, tính tương tác, khả năng truyền tải và lưu trữ lượng thông tin khổng lồ.

### ***a) Tính liên kết cộng đồng***

Đây là đặc điểm nổi bật của mạng xã hội ảo cho phép mở rộng phạm vi kết nối giữa người với người trong không gian đa dạng. Người sử dụng có thể liên kết với nhau, việc liên kết này tạo ra một cộng đồng mạng với số lượng thành viên lớn.

### ***b) Tính đa phương tiện***

MXH có rất nhiều tiện ích nhờ sự kết hợp giữa các yếu tố chữ viết, âm thanh, hình ảnh. Sau khi đăng kí tài khoản người sử dụng có thể tạo ra một không gian riêng cho bản thân. Nhờ các tiện ích đa phương tiện mà người sử dụng có thể chia sẻ thông tin, hình ảnh, video. Đặc điểm này được phản ánh trong cấu trúc phân lớp ứng dụng của MXH.

### ***c) Tính tương tác***

Tính tương tác được thể hiện không chỉ ở thông tin được di truyền và sau đó nhận được phản hồi từ phía người nhận mà còn phụ thuộc vào cách người dùng sử dụng các ứng dụng.

### ***d) Khả năng truyền tải và lưu trữ lượng thông tin khổng lồ***

Tất cả các MXH đều có những ứng dụng tương tự nhau như đăng trạng thái, nhạc, video clip, viết bài nhưng được phân bố với dung lượng khác nhau. Các trang MXH lưu trữ thông tin và nhóm sắp xếp chúng theo một thứ tự thời gian, nhờ đó người sử dụng có thể truy cập và tìm kiếm thông tin.

## **1.2. Lợi ích và các tác động của MXH**

Kể từ khi có kết nối mạng trên toàn cầu (Internet) và nhất là sau khi điện thoại thông minh hay máy tính bảng được chế tạo, việc sử dụng các MXH như Facebook, Instagram, Viber, Zalo, Skype, Whatsapp, Youtube, Linked, Twitter đã không còn xa lạ với hầu hết người dùng, kể cả trẻ em, thanh thiếu niên và người lớn tuổi. Thế giới ngày càng phát triển, mạng xã hội càng giúp con người xích lại gần nhau hơn, đem lại những tính năng và lợi ích vô cùng tuyệt vời.

### ***a) Kết nối bạn bè, gia đình, cộng đồng***

Ngày nay, con người ngày càng có ít thời gian cho bản thân và mở rộng các mối quan hệ. Nhờ có MXH, thông qua văn bản, video, hình ảnh con người có thể kết nối với nhau rất thuận tiện. Ngoài ra họ cũng có thể mở rộng các mối quan hệ khác về mọi lĩnh vực mà người dùng quan tâm. Người dùng có thể kết bạn với nhiều nhóm người với những sở thích, sở trường khác nhau. Hầu hết các MXH đều yêu cầu người dùng để đưa ra một số thông tin nhất định thường bao gồm: Độ tuổi, giới tính, địa điểm, quan điểm,

sở thích... Tuy nhiên, những thông tin rất cá nhân thường không được khuyến khích vì lý do an toàn. Điều này cho phép người dùng khác tìm kiếm theo một số loại tiêu chuẩn phù hợp đối với mình và duy trì một mức độ ẩn danh tương tự như hầu hết các dịch vụ hẹn hò trực tuyến.

#### ***b) Cập nhật tin tức, kiến thức, xu thế***

Tin tức sẽ được cập nhật theo từng giây, trong học tập, nghiên cứu thì đây cũng là một kênh tin tức bổ ích. Theo báo cáo của Hội Liên hiệp giáo dục Mỹ (The National School Boards Association), 60% sinh viên sử dụng mạng xã hội nói chuyện về chủ đề giáo dục trực tuyến, và hơn 50% nói chuyện cụ thể về việc học ở trường. Một số mạng xã hội khác như: TermWiki, Learn Central và các trang web khác được xây dựng để thúc đẩy mối các quan hệ trong giáo dục bao gồm các Blog giáo dục, ePortfolios cũng như thông tin liên lạc như chat, bài thảo luận, và các diễn đàn học tập.

#### ***c) Cải thiện chất lượng và tốc độ của báo chí và dịch vụ công***

Do tính năng cập nhật và lan rộng nhanh của MXH mà các cơ quan báo chí và thông tin đại chúng đang tích cực đăng tải cùng một lúc trên báo giấy, trên báo điện tử và trang mạng của mình để theo kịp xu thế của thời đại và giữ số lượng độc giả của mình. Các cơ quan pháp luật hay dịch vụ công cũng đang dần “lên sóng” MXH để cập nhật những tin tức và quy định mới của mình hoặc lắng nghe ý kiến phê bình góp ý của người dân nhằm giúp giảm thiểu sự quan liêu, phức tạp hay sai sót trong dịch vụ công, để tiến tới một bộ máy hành chính công thông minh và giản tiện hơn.

#### ***d) Cải thiện kĩ năng sống, kiến thức sống***

Hiện nay trên các MXH xuất hiện ngày càng nhiều các trang dạy ngoại ngữ, nấu ăn, sửa chữa, giao tiếp, tâm lý, thể thao để xem tham khảo, tự học mà không cần đến lớp hay đóng lệ phí. Chính nhờ tham gia các cộng đồng mạng này, chúng ta đang ngày càng trở nên hoàn thiện hơn với những kỹ năng cơ bản cần thiết trong cuộc sống hiện đại như sử dụng ngoại ngữ, cách giao tiếp văn minh hay có một thể hình khỏe đẹp.

#### ***e) Kinh doanh, quảng cáo miễn phí***

Rất nhiều công ty, nhà quảng cáo đã sử dụng MXH để bán hàng, quảng cáo cho sản phẩm của mình. Ở các trang mạng xã hội như Facebook, Instagram,... có rất nhiều người trẻ khởi nghiệp bằng cách bán hàng online. MXH kết nối con người với chi phí thấp, có những chức năng phù hợp, do đó đem lại hiệu quả cao, kể cả việc tư vấn cho người dùng cũng trở nên nhanh chóng, dễ dàng. Việc kinh doanh, quảng cáo sẽ trở nên phổ biến hơn do chức năng “chia sẻ” từ đó sẽ có nhiều hơn những người dùng đọc được



thông tin về sản phẩm.

***f) Tiết kiệm kinh phí, thời gian, sức lao động***

Nhờ MXH mà công ty, tổ chức hay hộ gia đình đã tiết kiệm được chi phí giấy, mực in, nhân công, phí điện thoại, tin nhắn. Một số mạng xã hội nhằm mục đích khuyến khích lối sống lành mạnh đối với người dùng. Ví dụ: MXH SparkPeople cung cấp cho cộng đồng các công cụ trợ đồng đẳng trong việc giảm cân, Fitocracy tập trung vào hướng dẫn người dùng trong tập thể dục hoặc cho phép người dùng chia sẻ tập luyện của mình và nhận xét về những người dùng khác.

***g) Tác động chính trị, xã hội***

Nếu được sử dụng đúng cách, MXH có vai trò quan trọng đối với các chính trị gia, giúp họ được nhiều người dân biết đến hơn, giúp họ tự xây dựng hình ảnh trong mắt công chúng. Từ đó, có thể thúc đẩy sự nghiệp của họ. Ví dụ điển hình cho việc sử dụng mạng xã hội thành công là Tổng thống Mỹ Donald Trump. Gần đây, ông Donald Trump đã khẳng định: “Tôi nghĩ có thể tôi đã không ngồi ở vị trí này nếu không có Twitter. Twitter là một thứ tuyệt vời đối với tôi, vì tôi có thể truyền tải suy nghĩ của mình tới công chúng. Tôi có thể sẽ không ở đây nói chuyện với tư cách tổng thống nếu không thể đưa ra những phát biểu chân thật”.

Dù vậy, MXH cũng là “con dao hai lưỡi” nếu ta sử dụng không đúng mục đích. Sử dụng MXH quá nhiều sẽ dẫn đến xao nhãng trong học tập, mất đi thời gian vận động, thể dục thể thao. Do nguồn thông tin trên mạng không có ai giám sát, kiểm duyệt nên còn tràn lan rất nhiều thông tin sai lệch, văn hóa phẩm đồi trụy, trong khi giới trẻ còn chưa đủ nhận thức để sàng lọc thông tin, dễ dẫn đến nhận thức lệch lạc, kéo theo đó là hành động sai lầm như: giết người, nghiện hút, mại dâm... Bên cạnh đó sử dụng MXH quá nhiều còn dẫn đến mất khả năng tương tác giữa mọi người, có nguy cơ mắc bệnh trầm cảm, xao nhãng những mục tiêu thật của cuộc sống. Việc bảo mật thông tin chưa thực sự tốt, nghiêm trọng nhất là nguy cơ lây lan thông tin, đặc biệt là những thông tin sai lệch trên MXH, sẽ dẫn đến những hậu quả vô cùng nghiêm trọng.

**2. Sự lây lan và tác hại của thông tin sai lệch trên MXH**

Trong phần này, nhóm tác giả trình bày định nghĩa thông tin sai lệch, các tính chất của thông tin sai lệch, thực trạng phát tán của thông tin sai lệch trên thế giới và thực trạng ở Việt Nam, đi kèm theo đó là các vụ việc thực tế đã và đang xảy ra trong thời gian gần đây.

## 2.1. Định nghĩa thông tin sai lệch và tính chất

### a) Định nghĩa thông tin sai lệch

Trong thực tế trên MXH luôn tồn tại những thông tin lệch lạc, không lành mạnh gây ra ảnh hưởng tiêu cực đến người dùng bên cạnh những giá trị tích cực mà chúng mang lại.

Theo Karlova và Fisher, 2013 [9], thông tin sai lệch được hiểu là những thông tin giả mạo, không chính xác. Dựa trên mục đích của người lan truyền, thông tin sai lệch được phân thành hai loại:

- *Thông tin sai lệch lan truyền vô ý*: Thông tin sai lệch được tạo ra và lan truyền một cách vô ý, không có chủ đích. Mọi người có xu hướng giúp lan truyền những thông tin như vậy do niềm tin với bạn bè, người thân và ảnh hưởng của họ trên MXH.
- *Thông tin sai lệch lan truyền cố ý*: Đó là những tin đồn, tin tức giả mạo, sai lệch được tạo ra và lan truyền một cách cố ý bởi người dùng với mục đích, động cơ không trong sáng.

Ở Việt Nam, tại khoản 1, điều 5 Nghị định 72/2013/NĐ-CP ngày 15/7/2013 của Chính phủ đã có quy định chi tiết về việc quản lý, cung cấp, sử dụng dịch vụ Internet và thông tin trên mạng [8]. Trong đó có nhiều hành vi bị nghiêm cấm như lợi dụng việc cung cấp, sử dụng dịch vụ Internet và thông tin trên mạng nhằm mục đích chống lại Nhà nước Cộng hòa xã hội chủ nghĩa Việt Nam; gây phương hại đến an ninh quốc gia, trật tự an toàn xã hội; phá hoại khối đại đoàn kết dân tộc; tuyên truyền chiến tranh, khủng bố; gây hận thù, mâu thuẫn giữa các dân tộc, sắc tộc, tôn giáo (điểm a). Tuyên truyền, kích động bạo lực, dâm ô, đồi trụy, tội ác, tệ nạn xã hội, mê tín dị đoan, phá hoại thuần phong, mỹ tục của dân tộc (điểm b). Tiết lộ bí mật nhà nước, bí mật quân sự, an ninh, kinh tế, đối ngoại và những bí mật khác do pháp luật quy định (điểm c). Đưa thông tin xuyên tạc, vu khống, xúc phạm uy tín của tổ chức, danh dự và nhân phẩm của cá nhân (điểm d). Quảng cáo, tuyên truyền, mua bán hàng hóa, dịch vụ bị cấm; truyền bá tác phẩm báo chí, văn học, nghệ thuật, xuất bản phẩm bị cấm (điểm đ). Giả mạo tổ chức, cá nhân và phát tán thông tin giả mạo, thông tin sai sự thật xâm hại đến quyền và lợi ích hợp pháp của tổ chức, cá nhân (điểm e).

Như vậy, có thể thấy rằng, mặc dù có những định nghĩa khác nhau về thông tin sai lệch tuy nhiên những khái niệm có những điểm tương đồng giống nhau. Đó đều là những thông tin không đảm bảo tính chính xác hoặc thông tin giả mạo, xuyên tạc vắn đề, xuyên tạc nội dung v.v... gây ảnh hưởng xấu đến cá nhân và tổ chức, đồng thời mỗi quốc gia

có những quy định riêng về những hành vi bị cấm khi đưa thông tin lên mạng và đề được cụ thể hóa trong văn bản pháp luật.

Xuất phát từ những thực tế nêu trên, nhóm tác giả nhận thấy việc ngăn chặn, hạn chế kịp thời sự lan truyền của thông tin sai lệch trên MXH là vô cùng cấp thiết nhằm giảm thiểu tối đa những thiệt hại do chúng gây ra đối với người dùng, góp phần làm trong sạch môi trường mạng, nâng cao sự tin tưởng của người dùng đối với những thông tin trên MXH.

### ***b) Tính chất của thông tin sai lệch***

*Thông tin xuyên tạc có thể có tốc độ lây lan nhanh, dễ dàng:* Trong thời đại truyền thông 4.0, MXH trở thành một môi trường thích hợp để thông tin xuyên tạc lây lan nhanh chóng bởi những đặc điểm của nó. MXH là những website mở, nội dung được xây dựng hoàn toàn bởi các thành viên tham gia, hơn thế nữa, MXH có sự tham gia trực tiếp của nhiều cá nhân hay các chủ thể. Qua đó, các thông tin được dư luận và các cá nhân cung cấp có thể mập mờ, không chính xác là cơ sở cho việc phát sinh thông tin xuyên tạc, sai lệch. Hơn thế nữa, người đọc thường có xu hướng quan tâm đến các tin đồn nhiều hơn tin chính thống vì chúng có thể gây nên nhiều liên tưởng tò mò, hấp dẫn. MXH Facebook cho phép người dùng đăng tin lên cùng với các chức năng bình luận, like, share. Chính vì vậy, thông tin sai lệch thông qua những đặc điểm này có thể lan truyền với tốc độ chóng mặt và ngày càng nhiều diễn biến phức tạp.

*Khả năng lan truyền rộng, khó kiểm soát được tính xác minh của thông tin:* Ban đầu thông tin sai lệch được đăng tải lên bởi một cá nhân hoặc tổ chức nhưng chưa được xác minh hoặc ghi nhận bởi các cơ quan thẩm quyền, thông tin này thường được xuất phát từ các nhóm nhỏ và thường mang nội dung về các vấn đề chính trị, về cá nhân tổ chức khác, chúng được thảo luận, chia sẻ qua các tính năng của mạng xã hội đến nhiều người dùng để tạo cơ sở để công chúng tin tưởng hoặc bị ảnh hưởng. Khi thông tin đó đã trở nên phổ biến, có được sự quan tâm rộng rãi của công chúng, trong quá trình truyền từ người này sang người khác, các đối tượng xấu đã bóp méo dần sự thật, thêm thắt gây ra hậu quả vô cùng to lớn.

Thế nhưng, hiện nay các mạng xã hội vẫn chưa có một cơ chế để kiểm duyệt những thông tin này có phù hợp không, người dùng cũng thiếu các thông tin xác thực để kiểm chứng, thậm chí dẫn đến hiểu nhầm, hiểu sai về bản chất của vấn đề.

## **2.2. Thực trạng sự phát tán của thông tin sai lệch trên mạng xã hội**

Không chỉ ở Việt Nam mà sự phát tán diện rộng của thông tin trên toàn thế giới đã trở thành một vấn nạn chưa từng có. MXH trở thành một trong những ưu tiên hàng đầu của các cơ quan hành chính cũng như tư nhân áp dụng vào nhiều mục đích khác nhau.

Tuy nhiên, điều này đi kèm với những nguy cơ quá lạm dụng các trang MXH.

MXH ảnh hưởng rất nhiều đến các hoạt động và các mối quan hệ trong thế giới thực. Trong đó, những tin tức về giải trí được quan tâm nhất. Khi người dùng đọc các tin tức mà họ quan tâm, họ có nhiều khả năng sẽ duy trì thảo luận quanh thông tin đó. Ngoài ra, khi nội dung thông tin liên quan đến vấn đề chính trị, người dùng có nhiều khả năng đưa ra quan điểm, ý kiến của mình về chính trị.

#### ***a) Thực trạng thế giới***

Với số người sử dụng các mạng xã hội trên toàn cầu vào khoảng 3 tỷ người và không có dấu hiệu dừng lại như hiện nay, thông tin sai lệch có sức ảnh hưởng vô cùng lớn tới tình hình thế giới theo nhiều khía cạnh. Sự lây lan thông tin độc hại ảnh hưởng mạnh đến tình hình kinh tế - chính trị, không những thế các đối tượng xấu sử dụng thông tin sai lệch để lừa đảo, chiếm đoạt ảnh hưởng đến tâm lý cũng như sức khỏe của người dùng.

Ngày 23 tháng 4 năm 2013 tin tặc giả mạo hãng thông tấn Associated Press tung tin Nhà Trắng bị đánh bom và cựu Tổng thống Obama bị thương nặng bởi một vụ nổ ở Nhà trắng. Ngay lập tức thông tin này làm thị trường tài chính chứng khoán của Mỹ chao đảo. Các chỉ số chứng khoán gần như sụp đổ bởi thông tin này. Chỉ số Dow Jones ngay lập tức sụt giảm đến 143 điểm gây thiệt hại 136,5 tỷ USD cho thị trường. Mặc dù vậy, thị trường chỉ rơi vào khoảng lặng hơn 1 phút trước khi AP thông báo đó là tin giả mạo do tài khoản Twitter của báo này bị tin tặc chiếm quyền điều khiển và đăng tin sai sự thật [10].

Gần đây những thông tin sai lệch trên các MXH còn được cho là có ảnh hưởng không nhỏ tới cuộc bầu cử ở Pháp và ở Mỹ. Trong cuộc bầu cử tổng thống ở Pháp, Facebook đã xóa 30.000 tài khoản giả mạo báo cáo tin đồn ở Pháp trước cuộc bầu cử Tổng thống vào năm 2017 [11]. Trong cuộc bầu cử ở Mỹ, MXH được cho ảnh hưởng không nhỏ tới kết quả bầu cử Tổng thống năm 2016. Nhiều tài khoản giả mạo được tạo ra chia sẻ những thông tin sai lệch về sơ hở rở rã email của bà Hillary Clinton và các đồng sự cấp cao của bà. Các tài khoản này đã phát động một chiến dịch phản đối bà Hillary đây được cho là một trong những nguyên nhân lớn đưa đến sự thất bại của bà.

Chính vì vậy, hiện nay các nước trên thế giới đã thành lập trung tâm chống tin giả, hiệp hội chống tin giả hay thông qua các luật an ninh mạng giúp việc ngăn chặn và phòng chống thông tin sai lệch ngày càng hiệu quả.

#### ***b) Thực trạng tại Việt Nam***

MXH trong những năm gần đây ngày càng trở nên thịnh hành tại Việt Nam. Cũng như trên thế giới, MXH (điển hình như Facebook) được nhiều người Việt Nam coi là tin

tưởng sử dụng. Chính vì vậy, các đối tượng luôn tìm cách lợi dụng điều này để có thể trục lợi cho bản thân, hay thực hiện những hoạt động chống phá, gây ảnh hưởng nghiêm trọng tới uy tín của cá nhân, tập thể, chính quyền.

Vào tháng 8 năm 2014, trên MXH lan truyền nhanh chóng nội dung tin đồn thất thiệt rằng “dịch Ebola đã bùng phát tại Hà Nội”. Trước đó, nhiều người đã truyền nhau thông tin tại Hà Nội đã có người nhiễm Ebola. Thông tin ngay sau đó đã được lan truyền một cách chóng mặt, gây nên một sự hoang mang lo lắng đến người dân. Bên cạnh đó còn có tin đồn cho rằng, nước láng giềng của Việt Nam là Campuchia đã xuất hiện ca nhiễm virus Ebola [12].

Tình trạng “ô nhiễm” thông tin xuất phát từ hoạt động phá hoại tư tưởng đồng thời phát tán các tư tưởng chống phá Đảng và nhà nước kích động biểu tình bạo loạn thông qua MXH của các thế lực thù địch, phản động chống đối ở trong và ngoài nước. Theo thống kê của cơ quan An ninh, tính đến nay có hơn 2500 trang web, blog, MXH của các cá nhân, tổ chức đang hoạt động, đăng tải các tin, bài viết, bình luận, bài phỏng vấn với mục đích tuyên truyền phá hoại tư tưởng nhằm phá hoại tư tưởng. Trong đó nổi lên một số trang như: danlambao, quanlambao, danluan. Các đối tượng quản trị những trang web này lợi dụng những điểm nóng về chính trị, xã hội và những thiếu sót trong công tác quản lý của ta để đăng tin xuyên tạc, kích động biểu tình, bạo loạn. Điển hình là hai vụ việc sự cố môi trường ở 4 tỉnh miền Trung do Formosa xả thải [13] và vụ việc khiếu kiện đất đai ở Đồng Tâm mà đỉnh điểm là vụ bắt giữa 38 chiến sĩ Cảnh sát cơ động. Những thông tin sai sự thật này gây phức tạp thêm tình hình gây khó khăn cho công tác giải quyết của Chính quyền Nhà nước.

MXH nơi để mọi người cùng nhau chia sẻ mọi điều trong cuộc sống và đồng thời cũng là nơi để mọi người cập nhật những tin tức, những hình ảnh mới, kết nối với nhau. Nhưng cũng chính từ đây, các chuyện hư cấu, tin đồn, chuyện bóp méo được đăng tải để “câu like”, gây ảnh hưởng đến cộng đồng, dư luận.

Những chiêu thức, trò đùa ác ý với nhiều mục đích khác nhau đã làm cho nhiều người hoang mang và tỏ ra e ngại khi tiếp nhận các thông tin trên MXH. Và không chỉ làm hoang mang dư luận, những tin đồn thất thiệt này đã trực tiếp làm ảnh hưởng đến cuộc sống của những người vô tình trở thành nạn nhân. Chỉ bằng một cú click vào xem, chia sẻ, người dùng có thể rơi vào cái bẫy khiến tài khoản cá nhân bị nguy hiểm, bị đánh cắp thông tin. Trong những đường dẫn chứa thông tin sai lệch này còn có thể kèm virus hoặc những phần mềm gián điệp nhằm lấy cắp thông tin hay chiếm quyền kiểm soát máy.

## **Kết luận chương**

Sự ảnh hưởng rộng lớn của thông tin sai lệch đến người dùng cũng như các công ty, doanh nghiệp ngày càng tăng mạnh nên việc tổ chức hạn chế sự ảnh hưởng của các loại thông tin này đang trở nên cấp bách hơn bao giờ hết.

Hậu quả của thông tin sai lệch trên MXH là vô cùng nghiêm trọng về mọi mặt chính trị, kinh tế, xã hội. Bản thân của cá nhân, tổ chức bị tung tin sai sự thật phải gánh chịu hậu quả, phiền toái không đáng có, thậm chí là những thiệt hại nặng nề về kinh tế, danh dự, phẩm chất. Nguy hại hơn những thông tin sai lệch về chính trị, đường lối chính sách của Đảng và Nhà nước còn gây mất lòng tin của nhân dân vào bộ máy. Những tin này là “mồi dẫn” để các thế lực thù địch tập hợp, lôi kéo lực lượng trên không gian mạng, tổ chức các cuộc biểu tình, bạo loạn, gây mất an ninh trật tự. Ở khía cạnh khác, đường link chia sẻ các loại tin sốc, bịa đặt được các hacker sử dụng để phát tán mã độc là bàn đạp cho các cuộc tấn công APT (Tấn công mạng sử dụng công nghệ cao), lừa đảo trên không gian mạng.

Trước những thực trạng to lớn hiện nay, nhóm đã mạnh dạn nghiên cứu, thực nghiệm để có thể đưa ra một giải pháp tối ưu để có thể ngăn chặn tác hại của lan truyền thông tin sai lệch trên mạng xã hội được trình bày ở các chương sau.

## CHƯƠNG 2

### THÔNG TIN SAI LỆCH VÀ CƠ CHẾ LAN TRUYỀN THÔNG TIN SAI LỆCH

Chương 1 đã trình bày tổng quan về các tác hại mà thông tin sai lệch gây ra đối với người dùng MXH. Những hậu quả to lớn của thông tin sai lệch trên MXH có thể thấy rõ trên hai khía cạnh sau:

- Về tâm lý, chính trị: Những thông tin sai lệch ảnh hưởng tiêu cực đến tâm lý, đời sống tinh thần của người dùng khi chúng được phát tán trên mạng. Nó có thể ảnh hưởng đến tinh thần, thái độ, thậm chí cả chính trị của khu vực người dùng sinh sống.

- Về kinh tế, những thông tin sai lệch tiêu cực về sản phẩm của một doanh nghiệp ảnh hưởng xấu đến tài chính, giá bán, doanh thu, và thậm chí là thương hiệu của doanh nghiệp đó.

Những tác hại kể trên cho thấy việc đối phó với các thông tin sai lệch là vô cùng cấp bách. Việc phát hiện nguồn thông tin sai lệch là cơ sở cho các giải pháp ngăn chặn sự phát tán của chúng. Nguồn phát tán thông tin sai lệch có thể được phát hiện thông qua khảo sát người dùng hoặc các phương pháp khai phá dữ liệu.

Để có thể đưa ra giải pháp hiệu quả trong việc ngăn chặn sự lan truyền của thông tin sai lệch, trước tiên chúng ta phải hiểu được cơ chế thông tin sai lệch lan truyền trên MXH. Chương này phân tích quá trình lan truyền thông tin sai lệch dưới hai mô hình lan truyền: Mô hình bậc độc lập và mô hình ngưỡng tuyến tính, đây là hai mô hình phổ biến, được sử dụng rộng rãi trong các công trình nghiên cứu liên quan đến vấn đề lan truyền thông tin, lan truyền ảnh hưởng trên MXH.

#### 1. Mô hình lan truyền thông tin

Thông tin được phát tán trên các MXH từ người dùng này đến người dùng khác thông qua nhiều hoạt động đăng bài, chia sẻ, bình luận. Kempe [15] là người đầu tiên đưa ra các mô hình phát tán thông tin, trong đó đã đưa ra hai mô hình phát tán thông tin cơ bản là Mô hình tầng độc lập (Independent Cascade – IC) và Mô hình ngưỡng tuyến tính (Linear Threshold – LT). Hai mô hình này sau đó đã được sử dụng rộng rãi trong các bài toán liên quan đến lan truyền, phát tán thông tin.

Các mô hình phát tán thông tin là cơ sở cho việc nghiên cứu hạn chế thông tin sai lệch cũng như các tin đồn thất thiệt. Các nghiên cứu về chủ đề này những năm gần đây đều dựa trên hai mô hình IC, LT và các biến thể của chúng [16], [17], [18]. Theo đó, một mạng xã hội được biểu diễn bởi các thành phần như sau:

- $V$  là tập hợp các đỉnh của đồ thị,  $|V| = n$ , biểu diễn những người dùng tồn tại trong MXH.

- $E \subset V \times V$  hợp các cạnh của đồ thị,  $|E| = m$  gồm  $m$  cạnh có hướng, biểu diễn mối quan hệ giữa các cá nhân trong MXH.
- $w(u, v)$  là trọng số của cạnh  $(u, v)$  là một số thực dương biểu diễn cho các tần số tương tác, trao đổi giữa hai người dùng.  $w(u, v) = 0$  nếu giữa hai đỉnh  $u$  và  $v$  không tồn tại cạnh,  $w(u, v) > 0$  nếu giữa  $u$  và  $v$  tồn tại cạnh nối.

Do  $G$  là đồ thị có hướng nên cạnh  $(u, v)$  được gọi là cạnh đi ra từ  $u$ , cạnh  $(v, u)$  được gọi là cạnh đi vào đỉnh  $u$ . Ta ký hiệu  $N_{out}(u)$  và  $N_{in}(u)$  tương ứng là tập hợp các đỉnh hàng xóm đi ra và đi vào đỉnh  $u$ .

Quá trình lan truyền thông tin theo các bước thời gian rời rạc, với thời gian  $t = 0, 1, 2, \dots$ . Gọi  $S_t \subset V$  là tập các đỉnh ở trạng thái *kích hoạt* tại thời điểm  $t$ . Tập các đỉnh là nguồn phát tán thông tin sai lệch ban đầu hay còn gọi là tập hạt giống, kí hiệu là  $S_0$

Tại mỗi bước thời gian  $t$ , đỉnh  $u$  ở trạng thái kích hoạt nếu  $u \in S_0$  hoặc  $u$  nhận được thông tin sai lệch từ các đỉnh hàng xóm ở trạng thái kích hoạt và chấp nhận thông tin này để tiếp tục chia sẻ, phát tán những thông tin sai lệch đó đến những đỉnh khác trong các bước tiếp theo, quá trình kích hoạt này ở mỗi mô hình lan truyền là khác nhau, ngược lại  $u$  ở trạng thái *không kích hoạt*.

Để mô tả mô hình lan truyền thông tin, các mô hình đang được nghiên cứu và sử dụng phổ biến, tiêu biểu trong số đó là: Mô hình ngưỡng (Threshold Model) [16], mô hình tầng (Cascading Model) [17], mô hình lan truyền ảnh hưởng cạnh tranh (Competitive Influence Diffusion Model) [18], mô hình dịch bệnh (Epidemic Model) [19]. Trong nội dung của đề tài, nhóm tác giả sử dụng mô hình ngưỡng tuyến tính (Linear Threshold – LT) và mô hình tầng độc lập (Independent Cascade – IC) [15] để mô tả quá trình lan truyền thông tin sai lệch trên MXH.

Trong nghiên cứu này, các tác giả đã tìm hiểu các mô hình trên qua đó có thể làm rõ tính chất về hành vi lan truyền thông tin của người dùng và chọn ra mô hình phù hợp để áp dụng và cũng đưa ra các mô hình phù hợp hơn với thực tiễn. Sau đây là sự mô tả chi tiết mô hình tiêu biểu được nhiều nghiên cứu sử dụng đó là: Mô hình tầng độc lập (IC) và Mô hình ngưỡng tuyến tính (LT).

- **Mô hình tầng độc lập:** Mô hình tầng độc lập (Independent cascade – IC) được đề xuất bởi Kempe [15] dựa trên các mô hình tương tác trong các hệ thống hạt và nghiên cứu về tiếp thị. Mô hình IC có liên quan tới mô hình dịch bệnh (Epidemic models). Đặc trưng chính của mô hình IC là quá trình lan truyền thông tin dọc theo các cạnh của đồ thị một cách độc lập với nhau.

- **Mô hình ngưỡng tuyến tính:** Mô hình IC phù hợp để mô tả sự lan truyền thông tin đơn giản, ở đó một đỉnh có thể được kích hoạt từ một đỉnh duy nhất. Tuy nhiên trong



thực tế có nhiều trường hợp một cá nhân cần nhiều sự tác động của các cá nhân khác để thay đổi hành vi của mình. Có thể kể đến các trường hợp như khi người dùng tiếp nhận một thông tin mới, một công nghệ mới, hay một thông tin sai lệch bôi xấu danh dự của các đồng chí lãnh đạo Đảng và Nhà nước, người dùng MXH cần được củng cố tích cực từ nhiều nguồn độc lập trong số bạn bè và người quen của họ trước khi họ thay đổi suy nghĩ của mình, chấp nhận thông tin đó. Các nhà khoa học đã đề xuất khái niệm *hành vi ngưỡng* để mô tả các hành vi kiểu như trên. Khi một hàm tổng hợp của các người dùng đã kích hoạt trên mạng đạt đến một ngưỡng nhất định thì đối tượng sẽ được kích hoạt, xét đến hành vi ngưỡng mà mỗi cá nhân chỉ được kích hoạt khi tiếp nhận ảnh hưởng từ nhiều hơn hai nguồn thông tin.

## 2. Cơ chế lan truyền thông tin sai lệch trên hai mô hình thông tin

### 2.1. Mô hình tăng độc lập

Trong mô hình IC, mỗi cạnh  $(u, v) \in E$  được gán một xác suất ảnh hưởng (Influence Probability)  $p_{uv} \in [0, 1]$  biểu diễn mức độ ảnh hưởng của đỉnh  $u$  với đỉnh  $v$ . Nếu  $(u, v) \notin E$ , thì  $p_{uv} = 0$ .

Quá trình lan truyền thông tin trên mô hình IC diễn ra theo bước thời gian rời rạc, tạo ra tập các đỉnh *kích hoạt* theo quy tắc sau:

- Tại thời điểm  $t = 0$ , tập đỉnh ở trạng thái kích hoạt chính là tập nguồn phát thông tin sai lệch  $S_0$ .
- Tại thời điểm  $t = 1$ , đầu tiên ta gán  $S_t$  bằng  $S_{t-1}$  sau đó với mỗi nút  $v \notin S_{t-1}$ , và với mỗi nút  $u \in N_{in}(v) \cap (S_{t-1} - S_{t-2})$ ,  $u$  thực hiện một lần thử kích hoạt bằng cách áp dụng phép thử Bernoulli (Phép tung đồng xu độc lập) với xác suất thành công là  $p(u, v)$ . Nếu thành công ta thêm  $v$  vào tập  $S_t$  và nói rằng  $u$  kích hoạt  $v$  tại thời điểm  $t$ . Nếu nhiều nút kích hoạt  $v$  tại thời điểm  $t$ , kết quả tương tự xảy ra,  $v$  được thêm vào tập  $S_t$ . Nói cách khác, sau khi nút  $u$  được kích hoạt tại thời điểm  $t - 1$ , ngay lập tức trong thời điểm  $t$ ,  $u$  có **một cơ hội duy nhất** để kích hoạt các hàng xóm chưa được kích hoạt  $v$  của nó với xác suất  $p(u, v)$ , và những sự kích hoạt này là độc lập với nhau. Nếu nút  $u$  không kích hoạt  $v$  tại thời điểm  $t$ , nó sẽ không thử kích hoạt  $v$  tại các thời điểm sau nữa. Và một khi một nút đã được kích hoạt, nó vẫn giữ trạng thái đó ở các bước sau.
- Nếu tại thời điểm  $t$ , không có nút nào được kích hoạt thêm nữa, nghĩa là  $S_t = S_{t-1}$ , tập các nút *kích hoạt* sẽ không còn thay đổi nữa, và quá trình truyền tin kết thúc với tập các nút bị kích hoạt cuối cùng là  $S_t$

Hình 2.1 chỉ ra một ví dụ của quá trình lan truyền thông tin trên mô hình IC. Các đỉnh màu da cam và màu xanh tương ứng biểu diễn các đỉnh ở trạng thái kích hoạt, và không kích hoạt. Cạnh liền màu đỏ từ  $u$  đến  $v$  biểu diễn  $u$  kích hoạt thành công  $v$ , cạnh nét đứt màu xanh từ  $u$  đến  $v$  biểu diễn  $u$  kích hoạt không thành công  $v$ .



Hình 2.1: Một số ví dụ quá trình lan truyền thông tin trên mô hình IC

Tại thời điểm bắt đầu  $t = 0$ , hai đỉnh  $v_1, v_2$  ở trạng thái *kích hoạt*. Ở thời điểm  $t = 1$ ,  $v_1$  kích hoạt thành công  $v_5$  nhưng thất bại với  $v_3$ , trong khi đó  $v_2$  kích hoạt thành công  $v_3$  và  $v_4$  nhưng thất bại với  $v_6$ . Tại thời điểm  $t = 2$ ,  $v_3$  kích hoạt thất bại  $v_6$  trong khi  $v_5$  kích hoạt thành công  $v_6$  nhưng thất bại với  $v_9$ . Ở bước  $t = 3$ ,  $v_6$  kích hoạt thất bại  $v_7$ , đến đây quá trình lan truyền thông tin kết thúc do không có đỉnh nào được kích hoạt thêm.

Mô hình IC phù hợp cho quá trình lan truyền thông tin, đó là các môi trường mà việc tiếp xúc với một nguồn là đủ để một cá nhân được kích hoạt.

## 2.2. Mô hình ngưỡng tuyến tính

Mô hình ngưỡng tuyến tính (Linear Threshold – LT) là mô hình khuếch tán ngẫu nhiên được đề xuất bởi Kempe [15]. Trong mô hình LT, mỗi cạnh  $(u, v) \in [0, 1]$  biểu diễn

mức độ ảnh hưởng của đỉnh  $u$  đến đỉnh  $v$ . Nếu  $(u, v) \notin E$  thì  $w(u, v) = 0$ . Các trọng số này được chuẩn hóa sao cho với mỗi đỉnh  $v$ , tổng trọng số tất cả các cạnh đi đến đỉnh  $v$  lớn nhất bằng 1, tức là:  $\sum_{u \in N_{in}(v)} w(u, v) \leq 1$ .

Tùy vào đặc tính của từng người dùng tương ứng, mỗi đỉnh  $v \in V$  có một giá trị  $\theta_v \in [0, 1]$ , biểu diễn ngưỡng đỉnh  $v$  bị ảnh hưởng bởi các đỉnh kích hoạt hàng xóm. Quá trình lan truyền thông tin trên mô hình LT diễn ra theo bước thời gian rời rạc, tạo ra tập các đỉnh kích hoạt theo quy tắc sau:

- Tại thời điểm  $t = 0$ , tập đỉnh ở trạng thái kích hoạt chính là tập nguồn phát thông tin sai lệch  $S_0$ .
- Tại thời điểm  $t \geq 1$ , đầu tiên ta gán  $S_t$  bằng  $S_{t-1}$ . Sau đó với mỗi đỉnh chưa được kích hoạt  $v \in S_{t-1}$ , nếu tổng ảnh hưởng từ những đỉnh hàng xóm kích hoạt tới  $v$  vượt ngưỡng  $\theta_v$ , tức là  $\sum_{u \in S_{t-1} \cap N_{in}(v)} w(u, v) \geq \theta_v$  thì đỉnh  $v$  được kích hoạt, ta thêm  $v$  vào tập  $S_t$ .
- Nếu tại thời điểm  $t$ , không có nút nào được kích hoạt thêm nữa, nghĩa là  $S_t = S_{t-1}$ , tập các nút kích hoạt sẽ không còn thay đổi nữa, và quá trình truyền tin kết thúc với tập các nút bị kích hoạt cuối cùng là  $S_t$ .

Sự ngẫu nhiên trong việc lựa chọn ngưỡng  $\theta_0$  từ 0 đến 1 phản ánh sự thiếu thông tin về ngưỡng nội bộ của mỗi cá nhân. Điều này phản ánh khá đúng với thực tế xã hội, bởi vì sự chấp nhận thông tin của mỗi người, tại những thời điểm khác nhau là khác nhau, và rất khó để nắm bắt.

Hình 2.2 chỉ ra một ví dụ quá trình lan truyền thông tin trên mô hình LT. Các đỉnh màu da cam và màu xanh tương ứng biểu diễn các đỉnh ở trạng thái *kích hoạt*, và *không kích hoạt*. Các cạnh liên màu đỏ cùng đến đỉnh  $v$  biểu diễn các cạnh này đồng thời cổ vũ kích hoạt đỉnh  $v$  và thành công.



Hình 2.2: Ví dụ quá trình lan truyền thông tin trên mô hình LT

Tại thời điểm  $t=0$ , tất cả các đỉnh được khởi tạo ngẫu nhiên ngưỡng  $\theta_0 \in [0, 1]$ , hai đỉnh  $v_1$  và  $v_2$  là các đỉnh hạt giống. Ở thời điểm  $t = 1$ ,  $v_1$  và  $v_2$  kích hoạt thành công  $v_3$ ,  $v_1$  cũng phải kích hoạt thành công  $v_5$  và  $v_2$  kích hoạt thành công  $v_4$ ; tuy nhiên  $v_6$  lại không kích hoạt thành công vì tổng trọng số các cạnh đi đến  $v_6$  là 0.3, trong khi ngưỡng kích hoạt của  $v_6$  0.7. Tại thời điểm  $t = 2$ , các đỉnh hàng xóm đi đến  $v_6$  là  $v_2$ ,  $v_3$ ,  $v_5$  đã được kích hoạt cho nên tổng trọng số các cạnh đi đến là 0.7 đủ để kích hoạt  $v_6$ . Tại bước  $t = 3$ , quá trình lan truyền thông tin kết thúc do không có đỉnh nào được kích hoạt thêm.

### CHƯƠNG 3

## GIẢI PHÁP NGĂN CHẶN SỰ PHÁT TÁN CỦA THÔNG TIN SAI LỆCH

Tự do ngôn luận, tự do báo chí là một trong những quyền căn bản của công dân đã được Hiến pháp ghi nhận và đã được Đảng và Nhà nước nhất quán. Tuy nhiên thực tế hiện nay có rất nhiều cá nhân, tổ chức đã và đang lợi dụng các quyền này để xâm phạm lợi ích của Nhà nước và lợi ích chính đáng của công dân. Với sự phát triển mạnh mẽ của Internet và mạng xã hội trực tuyến, việc đăng tải, tuyên truyền các thông tin sai lệch mang tính chất xuyên tạc, chống đối đường lối chính sách của Đảng, pháp luật của Nhà nước, hoạt động của bộ máy chính quyền diễn ra thường xuyên hơn với âm mưu, quy mô và tổ chức ngày càng rộng lớn, chặt chẽ.

Như thường lệ, mỗi khi cơ quan chức năng Việt Nam tiến hành truy tố, xét xử những cá nhân có hành vi vi phạm pháp luật như: Hoạt động nhằm lật đổ chính quyền nhân dân (theo Điều 109 Bộ luật Hình sự năm 2015, sửa đổi bổ sung năm 2017); tuyên truyền chống Nhà nước Cộng hòa XHCN Việt Nam (Điều 117) hay lợi dụng các quyền tự do, dân chủ, xâm phạm lợi ích của Nhà nước, quyền, lợi ích hợp pháp của tổ chức, công dân (Điều 331) thì ngay lập tức một số tổ chức, cá nhân chống đối trong và ngoài nước lại ráo riết tuyên truyền xuyên tạc, chống phá Việt Nam trên lĩnh vực dân chủ, nhân quyền. Những đối tượng này đã lợi dụng quyền tự do ngôn luận để viết và đăng tải các nội dung sai lệch, không có căn cứ, tuyên truyền xuyên tạc đường lối, chính sách của Đảng, pháp luật của Nhà nước, bôi nhọ các cán bộ, Đảng viên làm ảnh hưởng đến uy tín của cơ quan, tổ chức, đi ngược lại với lợi ích quốc gia, dân tộc.

Xuất phát từ những thực tế nêu trên, nhóm tác giả nhận thấy việc nghiên cứu đề ra giải pháp ngăn chặn kịp thời sự lan truyền của thông tin sai lệch trên MXH là một thách thức lớn cần giải quyết kịp thời. Từ thực trạng trên nhóm tác giả đề xuất xây dựng hai bài toán ngăn chặn.

Bài toán đầu tiên với nguồn ngân sách hạn chế, hạn chế bước thời gian lan truyền thông tin, yêu cầu tìm ra  $k$  đỉnh có vai trò quan trọng nhất đối với quá trình lan truyền thông tin. Bài toán thứ hai không hạn chế nguồn ngân sách, không hạn chế bước thời gian lan truyền thông tin, yêu cầu tìm tập ít đỉnh nhất đảm bảo số người dùng không bị nhiễm thông tin sai lệch vượt quá một ngưỡng cho trước.

Hai bài toán này đều dựa trên những yêu cầu thực tế, và bản chất của hai bài toán này là các bài toán tối ưu. Ứng với mỗi bài toán, nhóm tác giả đề xuất một thuật toán (giải pháp) tương ứng để giải quyết, đồng thời chứng minh độ tốt về mặt lý thuyết của thuật toán.

Bên cạnh đó, nhóm tác giả còn đưa ra các kết quả thực nghiệm của hai giải pháp

đã được đưa ra đối với hai bài toán tương ứng trên dữ liệu đã công bố, từ đó để thấy được hiệu quả của phương pháp nhóm đề xuất về kết quả và thời gian. Dữ liệu đã công bố là các bộ dữ liệu được lấy từ những nguồn uy tín hàng đầu, được các bài báo có chung mục tiêu nghiên cứu khác sử dụng và được các hội nghị, hội thảo thế giới công nhận. Việc thực nghiệm này sau khi đã kiểm chứng độ tốt về mặt lý thuyết cho thấy các giải pháp mà nhóm tác giả đề xuất có hàm lượng học thuật cao, khả thi về mặt lý thuyết.

### **1. Bài toán 1: Ngăn chặn thông tin sai lệch với ngân sách giới hạn**

Bài toán này được đề xuất và nghiên cứu sau khi xem xét thực tế quá trình lan truyền của một thông tin sai lệch trên mạng xã hội. Thông tin sai lệch luôn bắt đầu từ một số người dùng nhất định (có thể coi như đỉnh “nguồn” phát tán thông tin sai lệch). Những người dùng này thường là những kẻ có tư tưởng lệch lạc, phản động, và thông tin sai lệch thường là những thông tin bị bóp méo, những sai lầm của cán bộ bị thổi phồng, hoặc những lời lẽ bịa đặt vô căn cứ. Từ những người dùng bắt đầu này, thông tin sẽ được biết đến bởi những người xung quanh có liên kết với chúng, thường là những người dùng có quan hệ bạn bè với chúng. Thông tin này ban đầu có thể không tác động tới người dùng xung quanh, tuy nhiên nếu nó xuất hiện thêm nhiều lần từ nhiều nguồn khác nhau thì họ có thể sẽ tiếp nhận thông tin đó, và từ đó trở thành một nguồn phát thông tin sai lệch mới. Nếu mọi việc cứ phát triển thì thông tin này dần dần sẽ lan truyền sang toàn bộ một cộng đồng, hoặc thậm chí là toàn MXH (Quá trình lan truyền này được biểu diễn thông qua mô hình lan truyền được định nghĩa bên dưới). Do đó, cần có phương pháp giúp kiểm soát tình hình, ngăn chặn thông tin sai lệch lan truyền quá rộng, và đây cũng là lý do mà giải pháp này được nghiên cứu và công bố. Giải pháp được thiết kế để có thể chọn một số người dùng đặc biệt. Những người dùng được chọn này sau đó sẽ được các cơ quan chức năng có thẩm quyền tiến hành các phương pháp phù hợp để ngăn chặn khả năng phát tán thông tin của họ. Từ đó, thông tin sai lệch sẽ không có khả năng truyền thông qua những người dùng này, và dẫn đến khả năng lan truyền sang những người dùng tiếp theo bị cản trở. Giải pháp này đã được chứng minh rằng nó có thể lựa chọn ra các đỉnh hiệu quả nhất để ngăn chặn, do đó vừa đem lại hiệu quả cao mà không hao tổn quá nhiều tài nguyên vào việc ngăn chặn các đỉnh kém hiệu quả khác.

#### **1.1. Tổng quan bài toán**

Bài toán này được đưa ra nghiên cứu dựa trên những nhận định về việc thông tin lan truyền trên một mạng xã hội trong thực tế từ người dùng này đến người dùng khác có thể được coi như các bước lan truyền thông tin độc lập diễn ra liên tiếp nhau. Nếu ta càng ngăn chặn được sự lan truyền thông tin sai lệch sớm bao nhiêu, thì hậu quả mà thông tin sai lệch để lại sẽ càng được hạn chế bấy nhiêu. Không chỉ thế, các nghiên cứu

gần đây cũng đã chỉ ra rằng số các bước lan truyền thường nhỏ, ví dụ như trong [48] đã chỉ ra rằng số các bước lan truyền liên tiếp nhau thông thường chỉ nhỏ hơn 4, và trong [49] một bước lan truyền xảy ra thường là giữa hai người dùng có liên kết trực tiếp với nhau. Giải pháp tập trung vào việc tìm ra tập đỉnh  $A$  có tối đa  $k$  đỉnh để loại bỏ khỏi đồ thị để có thể tối thiểu hóa tác hại của lan truyền thông tin sai lệch. Những khía cạnh mới mà giải pháp của chúng tôi đề ra so với các công trình trước đó là: quá trình lan truyền được giới hạn trong khoảng  $d \geq 1$  bước lan truyền bắt đầu từ nguồn phát tán thông tin sai lệch, và giải pháp được thực hiện trên một mô hình lan truyền xác định bởi thực tế cho thấy thông tin cá nhân của người dùng có thể được thu thập qua khảo sát và các phương pháp khai phá dữ liệu [35]. Cụ thể, giải pháp này của chúng tôi có thể được tóm tắt lại như sau:

- Chúng tôi đưa ra một mô hình lan truyền thông tin với giới hạn về thời gian lan truyền thông tin, được mở rộng từ mô hình Ngưỡng tuyến tính xác định (Deterministic Linear-Threshold), gọi là mô hình Ngưỡng tuyến tính với bước thời gian rời rạc (Time Constraint Deterministic Linear-Threshold). Sau đó, chúng tôi định nghĩa bài toán Hạn chế sự lây lan của thông tin sai lệch (Limiting the Spread of Epidemics - LSE) với mục tiêu tìm kiếm tập đỉnh  $A$  có kích thước tối đa  $k$  để loại ra khỏi đồ thị sao cho số đỉnh cứu được là tối đa trong một ràng buộc thời gian. Chúng tôi chỉ ra bài toán này là bài toán NP-Khó.
- Với lời giải, chúng tôi đưa ra Thuật toán nhanh và hiệu quả để giới hạn sự lây nhiễm thông tin (Fast And Effective Limiting Epidemics – FLE).
- Việc kiểm thử được thực hiện trên các bộ dữ liệu thực tế lấy từ các nguồn đáng tin cậy như Gnutella, Wikipedia Vote, Amazon và Google Web, đồng thời chúng tôi cũng sử dụng một bộ dữ liệu được thu thập trực tiếp trên Facebook Việt Nam. Thuật toán đã được kiểm nghiệm và cho thấy khả năng ưu việt về cả tốc độ lẫn chất lượng lời giải so với các thuật toán phổ biến được dùng. Chi tiết về kiểm nghiệm sẽ được nêu rõ ở chương sau.

## **1.2. Mô hình lan truyền thông tin Ngưỡng tuyến tính với bước thời gian rời rạc (Time Constraint Deterministic Linear Threshold - T-DLT).**

Như đã được trình bày ở trên, có hai mô hình phổ biến trong bài toán lan truyền thông tin nói chung là mô hình IC và mô hình LT. Tuy nhiên, ở giải pháp này chúng tôi đề xuất một mở rộng mô hình DLT [35], vốn cũng là một trường hợp biến thể của mô hình LT, được gọi là T-DLT. Sở dĩ chúng tôi sử dụng mô hình này thay cho mô hình LT và IC bởi với bài toán này, mô hình T-DLT cho phép chúng tôi giảm thiểu tính phức tạp

của cơ chế lan truyền, đồng thời vẫn đảm bảo kết quả không bị ảnh hưởng so với hai mô hình điển hình trên. Chi tiết mô hình T-DLT có thể được mô tả như sau:

- *Ký pháp đồ thị*: Đặt  $G = (V, E, w)$  là biểu diễn của mạng xã hội với tập đỉnh  $V$ , tập cạnh có hướng  $E$ , với  $|V| = n$  và  $|E| = m$ . Mỗi đỉnh đại diện cho một người dùng trong mạng xã hội, mỗi cạnh  $e = (u, v)$  trong tập  $E$  tương ứng đại diện cho mối quan hệ giữa người dùng  $u$  và người dùng  $v$ . Chúng tôi ký hiệu đỉnh vào liền kề và đỉnh ra liền kề của đỉnh  $u$  lần lượt là  $N_{\text{in}}(u)$  và  $N_{\text{out}}(u)$ ,  $d_{\text{in}}(u)$  và  $d_{\text{out}}(u)$  lần lượt là bậc vào và bậc ra của đỉnh  $u$ . Đặt  $I \subset V$  là tập các đỉnh bị lây nhiễm ban đầu. Mỗi cạnh có hướng  $(u, v)$  sẽ có trọng số  $w(u, v)$ , tượng trưng cho mức độ ảnh hưởng thông tin của  $v$  tới  $u$ , thỏa mãn  $\sum_u w(u, v) \leq 1$ .
- *Trạng thái của đỉnh*: Quá trình phát tán thông tin sai lệch từ tập đỉnh ban đầu  $I$  tới các đỉnh còn lại trong mạng xã hội tiến triển theo từng bước thời gian rời rạc  $t = 1, 2, \dots, d$ . Mỗi đỉnh  $v \in V$  có hai trạng thái là *bị lây nhiễm* (hay “kích hoạt”) và *không bị lây nhiễm* (hay “không kích hoạt”).
- *Ngưỡng lây nhiễm* (hay *ngưỡng kích hoạt*): Mỗi đỉnh  $v$  sẽ có một ngưỡng lây nhiễm cho trước  $\theta_v \in [0, 1]$ . Giá trị này đại diện cho trọng số vào của đỉnh  $v$  cần chuyển thành bị lây nhiễm để  $v$  trở thành bị lây nhiễm.
- *Quá trình lan truyền*: Quá trình lây lan thông tin sai lệch được mô phỏng tuần tự theo từng bước thời gian rời rạc (còn gọi là “vòng”, “bước”)  $t = 0, 1, 2, \dots, d$ . Đặt  $I_t$  là tập các đỉnh bị lây nhiễm sau  $t$  bước, quá trình lan truyền được mô phỏng như sau:
  - + Tại bước  $t = 0$ , tất cả các đỉnh trong tập  $I$  đều bị lây nhiễm,  $I_0 = I$ .
  - + Tại bước  $t \geq 1$ , tất cả các đỉnh không bị lây nhiễm  $v$  sẽ chuyển thành bị lây nhiễm nếu tổng trọng số các cạnh vào gần kề chạm ngưỡng lây nhiễm của nó,  $\sum_{\text{các đỉnh vào gần kề của } u} w(u, v) \geq \theta_v$ .
  - + Một đỉnh bị lây nhiễm sẽ giữ nguyên trạng thái bị lây nhiễm cho tới hết quá trình lan truyền. Quá trình lan truyền chấm dứt khi  $t = d$ .

Trong mô hình LT, giá trị của ngưỡng  $\theta_v$  được đặt một cách ngẫu nhiên trong khoảng  $[0, 1]$  và sẽ được chỉnh sửa dựa vào thông tin có được thêm trong quá trình lan truyền. Do đó, mô hình này thuộc dạng mô hình ngẫu nhiên. Khác với mô hình LT, trong mô hình T-DLT ngưỡng lây nhiễm  $\theta_v$  được cho trước. Trường hợp này, giá trị ngưỡng có thể được đặt dựa theo các khảo sát thực tế hoặc các phương pháp khai phá dữ liệu.



Mô hình lan truyền T-DLT được chúng tôi lựa chọn để giải quyết bài toán bởi nó phù hợp với quá trình lan truyền thông tin nói chung trong thực tế. Khi một thông tin phát tán từ một nguồn, nó sẽ lan truyền từng bước thông qua những người có liên kết liên kề với nhau. Điều này được mô phỏng rõ dưới cách hoạt động của mô hình lan truyền T-DLT. Mặt khác, mô hình này có thể sử dụng các kết quả khai thác dữ liệu từ trước để thiết lập các thông số một cách linh động hơn hẳn so với các mô hình còn lại.

### 1.3. Định nghĩa bài toán

Đối với giải pháp này, việc ngăn chặn thông tin sai lệch được tiến hành bằng cách loại bỏ một số người dùng ra khỏi mạng để thông tin không thể truyền qua được, sao cho phạm vi lan truyền của thông tin là nhỏ nhất có thể. Nếu như coi mạng xã hội là một đồ thị, thì ta có thể hiểu rằng những đỉnh “bị lây nhiễm” là những người dùng đã bị ảnh hưởng bởi thông tin sai lệch và thông tin sai lệch đó có thể truyền từ người dùng này sang những người dùng thân cận, những đỉnh “không bị lây nhiễm” là những đỉnh không bị ảnh hưởng bởi thông tin sai lệch, còn những đỉnh “cứu được” là những đỉnh mà có trạng thái chuyển từ “bị lây nhiễm” sang “không bị lây nhiễm” khi chiến thuật loại bỏ đỉnh được áp dụng. Trong thực tế, việc loại bỏ một người dùng ra khỏi mạng xã hội có thể được thực hiện bằng cách thuyết phục người dùng đó thoát khỏi sự ảnh hưởng của thông tin sai lệch, hoặc ngăn chặn kết nối mạng của người dùng đó, hoặc xóa bỏ tài khoản của người dùng ra khỏi mạng xã hội, v.v. . . Tuy nhiên, những cách làm này đều là những cách làm yêu cầu có sự phối hợp từ nhiều bên liên quan, do đó nếu tiến hành loại bỏ quá nhiều sẽ dẫn đến chi phí phát sinh về thời gian và vật chất rất lớn, thậm chí sẽ gây ra trường hợp loại bỏ thành công như không kịp thời, khiến thông tin lan truyền quá rộng dẫn đến vô ích. Vì vậy, vấn đề đặt ra là phải làm sao chọn ra được một số lượng  $k$  đỉnh nào đó để khi thực hiện chiến thuật loại bỏ, ta có được số đỉnh cứu nhiều nhất. Số lượng  $k$  sẽ được tính toán một cách cân đối, cẩn thận nhất về chi phí để không làm mất tác dụng của chiến thuật loại bỏ. Chiến thuật loại bỏ  $k$  đỉnh này có thể được mô hình hóa và phát biểu khoa học như sau:

Kí hiệu  $f_d(I)$  là tập hợp các đỉnh đã bị lây nhiễm sau  $d$  vòng trên đồ thị  $G = (V, E)$  trong mô hình T-DLT,  $f_d(I, A)$  là tập các đỉnh bị lây nhiễm sau khi loại bỏ một tập các đỉnh  $A \subseteq V \setminus I$  từ  $G$  (số đỉnh bị lây nhiễm trong đồ thị sót lại). Khi đó, số đỉnh cứu được khi loại bỏ tập  $A$  là:  $h_d(A) = |f_d(I, \emptyset)| - |f_d(I, A)|$ .

Trong mô hình T-DLT, ta xây dựng một bài toán tối ưu tổ hợp là: *Hạn chế sự lây lan của thông tin sai lệch* (Limiting the Spread of Epidemics - LSE), có mục tiêu tìm kiếm một tập đỉnh có kích cỡ tối đa  $k$  đỉnh sao cho khi loại bỏ tập đỉnh đó, số đỉnh cứu được sẽ đạt cực đại.

**Định nghĩa 3.1** (Bài toán LSE). Cho đồ thị vô hướng  $G = (V, E)$  biểu diễn một mạng xã hội trong mô hình T-DLT. Một tập các đỉnh bị lây nhiễm ban đầu  $I \subset V$ , số vòng lan truyền  $d$  và số đỉnh tối đa loại bỏ được là  $k > 0$ . Tìm một tập  $k$  đỉnh  $A \subseteq V \setminus I$  sao cho khi loại bỏ khỏi mạng thì số đỉnh cứu được sau  $d$  vòng là  $h_d(A)$  đạt cực đại.

Kí hiệu  $G_d = (V_d, E_d)$  đồ thị con của đồ thị  $G = (V, E)$  trong đó  $V_d$  là tập các đỉnh mà khoảng cách giữa một đỉnh bất kì trong đó với một đỉnh trong tập  $I$  tối đa là  $d$ , và tập  $E_d$  là tập các cạnh nối giữa các đỉnh thuộc tập  $V_d$ , đồng thời  $n_d = |V_d|$ ,  $m_d = |E_d|$ . Ta thấy rằng sự lan truyền thông tin sai lệch chỉ xảy ra bên trong đồ thị con  $G_d$ . Do đó, để đơn giản hóa, thay vì xét đến tất cả các đỉnh thuộc  $G$ , ta sẽ chỉ tìm kiếm kết quả trong đồ thị con  $G_d$ .

#### 1.4. Độ phức tạp của bài toán LSE

Trong mục này, chúng tôi chỉ ra tính NP-Khó của bài toán LSE bằng phép dẫn đa thức về bài toán LSE từ bài toán Phủ đỉnh (Set Cover).

**Định lý 3.1.** *LSE là bài toán NP-Khó trong mô hình T-DLT.*

*Chứng minh.* Để chứng minh LSE là bài toán NP-Khó, ta sử dụng phép dẫn đa thức từ bài toán Phủ đỉnh (Set Cover) về bài toán LSE.

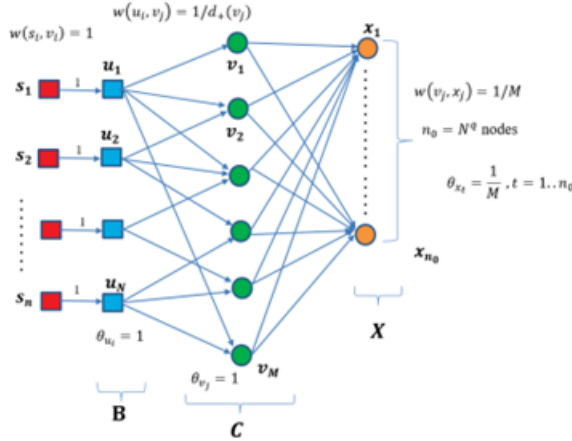
**Bài toán Phủ đỉnh (Set Cover - SC):** Cho một số tự nhiên dương  $t$ , một tập phổ quát  $\vartheta = e_1, e_2, \dots, e_M$  và tập con  $S = s_1, s_2, \dots, s_N$  ta có thể giả sử rằng  $t < M < N$ . Bài toán Phủ đỉnh yêu cầu rằng: có tồn tại hay không một lượng  $t$  tập con, sao cho hợp của chúng là  $\vartheta$ .

Bài toán SC đã được chứng minh là bài toán NP-Khó khi kích thước các tập được giới hạn là 3 và mỗi phần tử xuất hiện trong đúng hai tập con. Để giản thể bài toán SC về bài toán LSE, đầu tiên chúng tôi xây dựng một biểu diễn  $J_{LSE}$  của bài toán LSE từ biểu diễn  $J_{SC}$  của bài toán SC. Sau đó chúng tôi chỉ ra rằng nếu  $J_{SC}$  có tồn tại lời giải  $S$  kích thước  $t$ , thì  $J_{LSE}$  có lời giải  $A$ ,  $|A| \leq k$  với  $h_d(A) \geq k + M$  và ngược lại.

**Xây dựng phép dẫn:** Cho một biểu diễn  $J_{SC} = (\vartheta, S, t)$  của bài toán SC, ta xây dựng biểu diễn  $J_{LSE} = (G, I, d, k)$  của bài toán LSE như theo hình 3.1

- *Tập các đỉnh và các cạnh:* với mỗi  $S_i \in S$ , ta xây dựng một đỉnh bị lây nhiễm ban đầu  $s_i \in I$  và một đỉnh  $u_i$ . Ta có thêm một cạnh có hướng  $(s_i, u_i)$ . Với mỗi phần tử  $e_j \in \vartheta$  ta thêm một đỉnh  $v_j$  và thêm một cạnh có hướng  $(u_i, v_j)$  với mỗi  $u_i$ . Để thuận tiện, ta đặt  $B = u_1, u_2, \dots, u_N$ ,  $C = v_1, v_2, \dots, v_M$ .
- *Ngưỡng lây nhiễm và trọng số:* ta đặt  $w(u, v) = 1$ ,  
 $w(u_i, v_i) = \frac{1}{d_{in}(v_j)}$ ,  $\theta_{u_i} = \theta_{v_i} = 1$ .

- Sau cùng, ta cho  $k = t, d = 2$ .



Hình 3.1: Phép dẫn đa thức từ bài toán SC về bài toán LSE.

**Phân tích:** Qua bước xây dựng, ta thấy rằng  $|f_d(I, \emptyset)| = M + N$ . Nếu như có đỉnh bất kì  $u_i \in B$  là đỉnh kề vào và  $v_j \in C$  là đỉnh chưa bị lây nhiễm thì:

$$\sum_{\text{Các đỉnh hàng xóm tới đã bị lây nhiễm } u} w(u, v_j) \leq 1 - \frac{1}{d_{in}(v_j)} < 1 = \theta_{v_j}.$$

Do đó  $v_j$  là đỉnh không bị lây nhiễm. Nếu không, các đỉnh trong  $C$  là các đỉnh không bị lây nhiễm.

( $\rightarrow$ ) Giả sử  $S'$  là lời giải của biểu diễn  $J_{SC}$ , có nghĩa là  $|S'| = t = k$  và nó phủ  $t$  phần tử của  $\vartheta$ . Nếu ta chọn tập  $A$  chứa đỉnh  $u_i$  tương ứng với  $S_i \in S'$ , với mọi đỉnh  $v_i \in B$  kề với ít nhất một đỉnh trong  $A$ . Bằng bước phân tích trên, mọi đỉnh trong  $C$  là đỉnh không bị lây nhiễm. Ta có  $h_d(A) = t + M = k + M$ .

( $\leftarrow$ ) Ngược lại, nếu  $J_{LSE}$  có lời giải  $A$ ,  $|A| \leq k$  với  $h_d(I, A) \geq k + M$ . Nếu  $A$  chứa  $t_1$  ( $1 \leq t_1 \leq k$ ) đỉnh trong  $C$ ,  $h_d(I, A) \leq k - t_1 + M < k + M$ . Do đó  $A$  không chứa bất kì đỉnh nào thuộc  $C$ . Ở đây,  $A \subset B$ . Kết hợp với  $h_d(I, A) \geq k + M$ , mọi đỉnh thuộc  $C$  là đỉnh không bị lây nhiễm. Do đó, mọi đỉnh  $u_i \in A$  kề ít nhất với một đỉnh thuộc  $C$ . Vì vậy, qua bước xây dựng ta thấy rằng,  $S' = \{S_i | u_i \in A\}$  là lời giải cho bài toán  $J_{SC}$ . Định lý được chứng minh.  $\square$

### 1.5. Thuật toán đề xuất

Trong phần này, chúng tôi thiết kế một thuật toán heuristics có độ phức tạp thấp và kết quả rất khả quan. Tên gọi của thuật toán là Thuật toán nhanh và hiệu quả để giới hạn sự lây nhiễm thông tin (Fast And Effective Limiting Epidemics – FLE). Trước khi bắt đầu, chúng tôi có định nghĩa một vài kí hiệu sẽ được sử dụng trong quá trình đề xuất thuật toán:

- $t(u)$ : vòng nhỏ nhất sao cho đỉnh  $u$  chuyển từ đỉnh không bị lây nhiễm sang đỉnh bị lây nhiễm.
- $a_{\text{in}}(u) = \sum_{I_{t(u)-1} \cap N_{\text{in}}(u)} w(u, v)$  : tổng trọng số các cạnh vào từ các đỉnh vào hàng xóm đã bị lây nhiễm trước vòng  $t(u)$ .
- $\alpha(u) = \sum_{v \in U_{i=t(u)+1}^d} I_i \cap N(u) w(u, v), i = t(u) + 1, \dots, d$ : tổng trọng số các cạnh ra từ  $u$  tới các đỉnh bị lây nhiễm  $v$  tại vòng  $i$ .
- $\beta(u)$ : số lượng các đỉnh ra cứu được, sao cho chúng chuyển từ bị lây nhiễm thành không bị lây nhiễm sau khi loại bỏ đỉnh  $u$  ra khỏi đồ thị.

---

**Algorithm 1:** Fast Limit Epidemics (FLE) algorithm.

---

**Data:** Graph  $G = (V, E, \emptyset)$ , set of initial infected nodes  $I$ , propagation hop  $d$ .

**Result:** Set of nodes  $A$

```

1  $S \leftarrow \emptyset$ ;
2 Calculate  $G' \leftarrow G_d(I); U \leftarrow V_d$ 
3 for  $i = 1$  to  $k$  do
4   Calculate  $\alpha(u), \beta(u)$  on  $G'$  (Algorithm 3)
5    $u_{\text{max}} \leftarrow 0$ 
6   if  $\beta(v) = 0, \forall u \in U$  then
7      $u_{\text{max}} \leftarrow \arg \max_{v \in U} \alpha(v)$ 
8   else
9      $u_{\text{max}} \leftarrow \arg \max_{v \in U} \beta(v)$ 
10  end
11   $S \leftarrow S \cup \{u_{\text{max}}\}$ 
12   $U \leftarrow U \setminus \{u_{\text{max}}\}$ 
13  Remove  $u_{\text{max}}$  and all edges that adjacent with  $u_{\text{max}}$  from  $G'$ 
14 end
15 return  $S$ 

```

---

Một cách trực quan, số lượng gia tăng các đỉnh được cứu  $\delta(A, u)$  có thể xấp xỉ theo  $\beta(u)$ . Bên cạnh đó, để tăng cường tính hiệu quả của thuật toán đề xuất, chúng tôi kết hợp  $\alpha(u)$  và  $\beta(u)$  để tính toán khả năng lan truyền tin sai lệch của đỉnh  $u$ . Ý tưởng chính của thuật toán là chọn ra các đỉnh một cách lần lượt theo đánh giá của hai hàm  $\alpha$  và  $\beta$ . Ban đầu, chúng tôi khởi tạo tập đỉnh được chọn  $A = \emptyset$  và đặt tập các đỉnh được xét chọn  $U$  bằng với  $V_d$ . Trong mỗi bước, chúng tôi chọn ra đỉnh có trị số  $\beta(u)$  lớn nhất trong đồ thị còn lại. Trường hợp tất cả các đỉnh đều có giá trị  $\beta(u)$  là 0, chúng tôi chọn đỉnh có giá trị  $\alpha(u)$  cực đại. Chi tiết thuật toán được mô tả trong Thuật toán 1.

Khó khăn trong thuật toán chủ yếu đến từ bước tính toán các hàm  $\alpha$  và  $\beta$ . Để giải quyết việc này, chúng tôi bắt đầu với ý tưởng tìm kiếm theo chiều rộng (Breadth First Search - BFS) để tính toán  $f_d(I)$  và dần dần cập nhật hàm  $\alpha_{in}()$ . Sau đó, chúng tôi sử dụng nó để tính toán các hàm  $\alpha()$  và  $\beta()$  theo định nghĩa của chúng. Chi tiết thuật toán được thể hiện trong Thuật toán 2 và Thuật toán 3.

---

**Algorithm 2:** Calculate  $f_d(I), a_{in}(), t()$

---

**Data:** Graph  $G = (V, E, \emptyset)$ , set of infected nodes  $I$ , propagation hop  $d$ .

**Result:**  $f_d(I), a_{in}(u), t(u), \forall u \in G$ .

```

1  foreach  $u$  in  $G$  do
2      |    $t(u) \leftarrow +\infty; a_{in}(u) \leftarrow 0; \alpha(u) \leftarrow 0; \beta(u) \leftarrow 0$ 
3  end
4  foreach  $s$  in  $I$  do
5      |    $t(s) \leftarrow 0$ 
6  end
7  Queue  $Q \leftarrow \emptyset$ 
8  while  $Q \neq \emptyset$  do
9      |    $u \leftarrow Q.pop()$ 
10     |    $f_d(I) \leftarrow u$ 
11     |   if  $t(u) < d$  then
12         |   foreach  $v \in N_{out}(u)$  do
13             |   if  $t(u) < t(v)$  then
14                 |    $a_{in}(v) = a_{in}(v) + w(u, v)$ 
15                 |   if  $a_{in}(v) \geq \theta_v$  and  $t(u) + 1 < t(v)$  then
16                     |    $t(v) \leftarrow t(u) + 1$ 
17                     |    $Q.push(v)$ 
18                 |   end
19             |   end
20         |   end
21     |   end
22 end

```

---

Để đánh giá độ phức tạp của thuật toán này, chúng ta cần đánh giá trước độ phức tạp của thuật toán. Ta dễ thấy rằng việc tính toán  $f_d(I)$  (Thuật toán 2, dòng 7 - 22) có cách thức hoạt động giống BFS. Do đó, độ phức tạp của nó là  $O(m_d + n_d)$ . Với bước tính toán  $\alpha(u)$  và  $\beta(u)$  (Thuật toán 3, dòng 1 - 11), ta thấy độ phức tạp của nó là  $O(m_d + n_d)$ .

---

**Algorithm 3:** Calculate  $\alpha()$  and  $\beta()$ 

---

**Data:** Graph  $G = (V, E, \emptyset)$ , set of infected nodes  $I$ , propagation hop  $d$ ,  
 $f_d(I)$ ,  $a_{in}()$ ,  $t()$ .

**Result:**  $\alpha(u), \beta(u), \forall u \in G$ .

```

1 foreach  $u$  in  $f_d(I)$  do
2    $\alpha(u) \leftarrow 0; \beta(u) \leftarrow 0$ 
3   foreach  $v$  in  $N_{out}(u) \cap f_d(I)$  do
4     if  $t(u) < t(v)$  then
5        $\alpha(u) \leftarrow \alpha(u) + w(u, v)$ 
6       if  $a_{in}(v) - w(u, v) < \theta_v$  then
7          $\beta(u) \leftarrow \beta(u) + 1$ 
8       end
9     end
10  end
11 end

```

---

Do đó, độ phức tạp chung của thuật toán là  $O(m_d + n_d)$ .

### 1.6. Kết quả thực nghiệm với dữ liệu đã công bố

Trong phần này, chúng tôi thực nghiệm thuật toán đối với dữ liệu sẵn có trên một số mạng xã hội trực tuyến để kiểm chứng tính hiệu quả của thuật toán khi so sánh nó với các thuật toán cơ sở được sử dụng trong các bài toán lan truyền thông tin [27], [15], [30]. Chúng tôi so sánh đánh giá trên ba yếu tố chính: chất lượng lời giải, tính mở rộng, tác động của số vòng lan truyền  $d$  trên một vài các mạng xã hội trực tuyến thực tiễn. Các thuật toán cơ sở sẽ sử dụng bao gồm:

- Random: Thuật toán chọn ngẫu nhiên  $k$  đỉnh trong  $N_d(S)$ .
- Maxdegree (Maxdeg): Thuật toán heuristic chọn lần lượt các đỉnh có bậc cao nhất trong tập  $V_d$  cho đến khi đủ  $k$  đỉnh.
- Greedy: Thuật toán heuristic chọn lần lượt  $k$  đỉnh, trong mỗi lần chọn sẽ chọn ra đỉnh sao cho khi loại bỏ đỉnh đó sẽ cứu được nhiều đỉnh nhất có thể.

Chúng tôi thực nghiệm sử dụng một hệ thống máy có cấu hình sau: Intel(R) Core(TM) i5-6200U CPU @ 2.30 GHz (up to 2.40 GHz), bộ nhớ RAM 4GB, ngôn ngữ lập trình Python 2.7.

Mạng	Số đỉnh	Số cạnh	Loại	Bậc trung bình
Gnutella	6,301	20,777	Có hướng	3.29
Wikipedia vote	7,115	103,689	Có hướng	14.57
Amazon	262,111	1,234,877	Có hướng	4.71
Google web	875,713	5,105,039	Có hướng	5.83

Bảng 3.1: Các bộ dữ liệu

**a) Dữ liệu:**

Chúng tôi thực nghiệm thuật toán trên nhiều bộ dữ liệu thực tế khác nhau. Bên cạnh việc lựa chọn các bộ dữ liệu để đảm bảo sự đa dạng về kích thước, chúng tôi cũng lựa chọn nhiều miền dữ liệu khác nhau, bảng 3.1 thể hiện các bộ dữ liệu nhóm tác giả sử dụng.

**b) Thiết lập tham số**

Chúng tôi sử dụng các thiết lập tham số dưới đây cho việc đánh giá thực nghiệm:

- Trọng số của cạnh:  $w(u, v) = \frac{1}{d_{\text{in}}(v)}$ , với  $d_{\text{in}}(v)$  là bậc vào của đỉnh  $v$  [15] [40] [41] [47].
- Ngưỡng lây nhiễm: chúng tôi lựa chọn ngưỡng lây nhiễm trong tập  $\{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ .
- Số vòng lây nhiễm: chúng tôi lựa chọn số vòng lây nhiễm  $d = 2, 3, 4, 5$  dựa theo một nghiên cứu có trước [48].
- Số đỉnh bị nhiễm ban đầu: trong mọi bộ dữ liệu, chúng tôi lấy 1% số đỉnh làm tập đỉnh bị lây nhiễm ban đầu.

**c) Các thuật toán sử dụng**

Chúng tôi sử dụng các thuật toán sau đây cho việc đánh giá thực nghiệm cùng với thuật toán FLE, các thuật toán này đều là các thuật toán được sử dụng trong các nghiên cứu liên quan đã có trước:

- *Random*: Thuật toán lựa chọn ngẫu nhiên  $k$  đỉnh trong đồ thị mạng xã hội để loại bỏ.
- *Greedy*: Thuật toán thực hiện  $k$  lần, mỗi lần chạy thuật toán sẽ thử loại bỏ lần lượt từng đỉnh ra khỏi đồ thị và đánh giá số lượng đỉnh cứu được tương ứng, qua đó lựa chọn ra đỉnh loại đi mà cứu được nhiều đỉnh nhất để loại bỏ.

- *MaxDegree*: Thuật toán thực hiện  $k$  lần, mỗi lần chọn thuật toán sẽ chọn ra đỉnh có bậc lớn nhất để loại nó ra khỏi đồ thị.

#### ***d) Kết quả thực nghiệm***

##### *- Kết quả lời giải:*

Chúng tôi đo đặc hiệu suất của thuật toán trong ba trường hợp: (1) giá trị  $k$  thay đổi từ 10 đến 100,  $d = 5$ ,  $\theta = 0.5$  (hình 3.2); (2) ngưỡng  $\theta$  thay đổi,  $d$  và  $k$  giữ nguyên giá trị 50 (hình 3.3); (3) giá trị vòng lặp nhiệm  $d$  thay đổi,  $\theta$  và  $k$  không đổi (hình 3.4). Trong mọi trường hợp, FLE và Greedy cho kết quả tốt hơn so với các thuật toán còn lại, cụ thể là số lượng đỉnh cứu được nhiều hơn 48.5 lần so với thuật toán MaxDegree và Random trong hai bộ dữ liệu Gnutella và Wiki Vote. Khi so sánh Greedy và FLE, chúng tôi thấy Greedy có hiệu suất tốt hơn từ 1.02 đến 1.05 lần so với FLE trong bộ dữ liệu mạng Gnutella. Tuy nhiên, khoảng cách này bị thu hẹp lại khi  $\theta$  và  $k$  tăng. Đặc biệt khi  $k \geq 50$  và  $\theta \geq 0.4$ , hiệu suất của hai thuật toán là tương đương nhau. Hình 3.2, 3.3, 3.4 cho thấy Greedy và FLE có hiệu suất tương đương nhau trong bộ dữ liệu mạng Wiki Vote. Trong khi Greedy không thể chạy trong thời gian cho phép trên bộ Amazon và Google Web, FLE vẫn có thể đưa ra kết quả tốt hơn nhiều so với hai thuật toán còn lại.

##### *- Kết quả thời gian:*

Thời gian chạy của thuật toán cũng được mô tả trong hình 3.2, 3.3, 3.4. Đúng như dự đoán, thời gian chạy của Greedy là cực kỳ cao so với các thuật toán còn lại, chiếm 4.5 phút cho bộ dữ liệu Gnutella và 20.2 phút cho mạng Wiki Vote. Thuật toán FLE nhanh hơn từ 4820 đến 6789 lần so với Greedy trong mạng Gnutella và nhanh hơn từ 5839 đến 14490 lần so với Greedy trong mạng Wiki Vote. Điều này xảy ra do Greedy có độ phức tạp thuật toán lớn  $O(n_d k(m_d + n_d))$ , trong khi FLE chỉ có độ phức tạp nhỏ hơn nhiều  $O(k(m_d + n_d))$ . Trong bộ dữ liệu Amazon và Google Web, Greedy không thể tìm thấy lời giải trong vòng 12 tiếng và bị buộc dừng chạy, trong khi FLE chỉ mất 0.45 giây và 7.8 giây tương ứng với mỗi bộ dữ liệu trên. Điều này cho thấy FLE vẫn chạy tốt ngay cả với các bộ dữ liệu lớn.

##### *- Ảnh hưởng của tham số $d$ :*

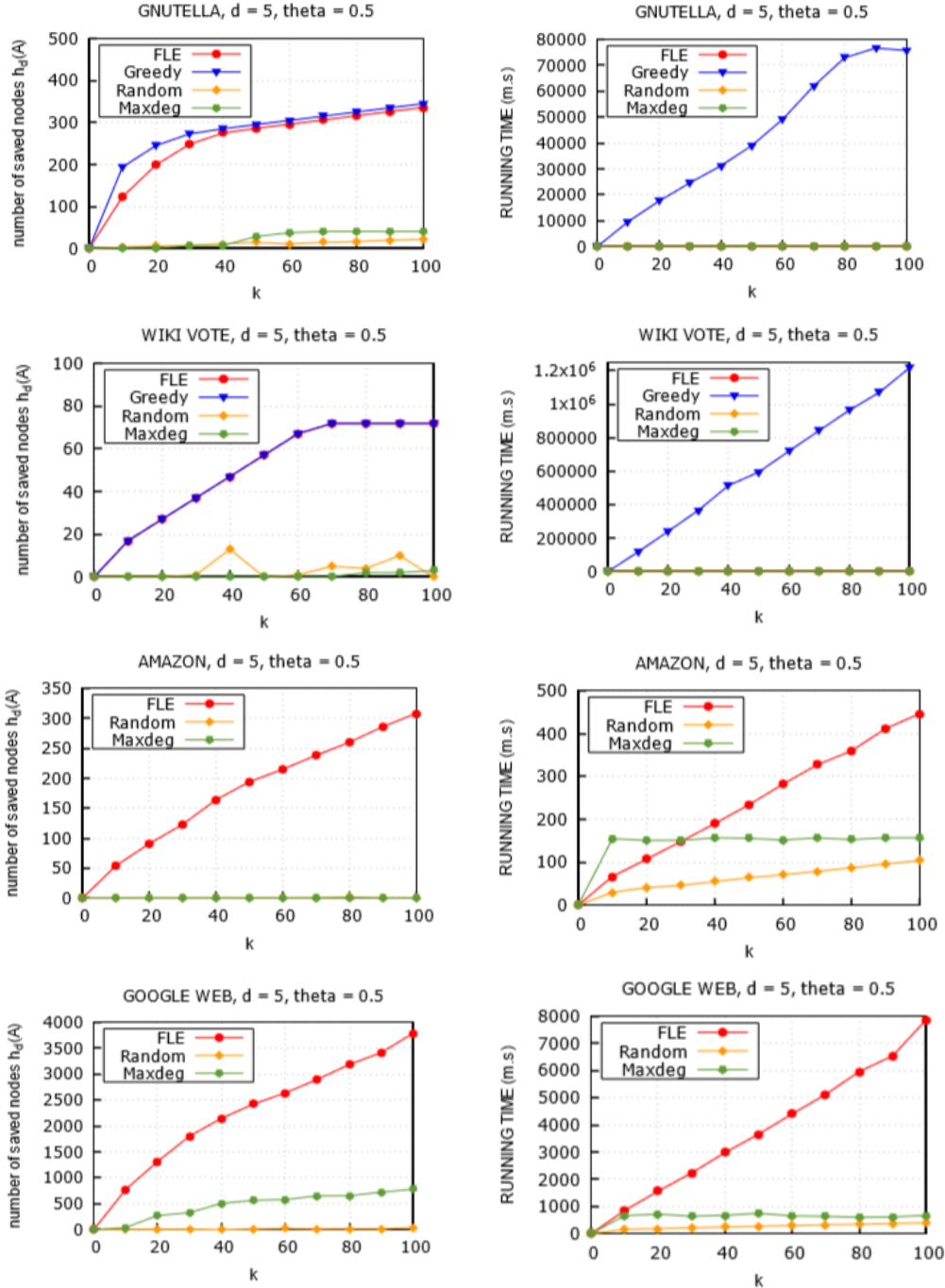
Chúng tôi khám phá ảnh hưởng của tham số  $d$  đối với các thuật toán khác nhau. Cho  $d$  thay đổi từ 2 đến 5 trong mạng Gnutella, kết quả được thể hiện trong hình 3.4. Với Greedy và FLE, số đỉnh cứu được tăng khi giá trị  $d$  tăng. Đặc biệt, số đỉnh cứu được tăng mạnh với  $d = 2, 3$  và tăng ít hơn với  $d = 4, 5$ . Điều này chứng tỏ để ngăn chặn lây lan, ta phải loại bỏ đỉnh càng sớm càng tốt.

##### *- Ảnh hưởng của tham số $\theta$ :*

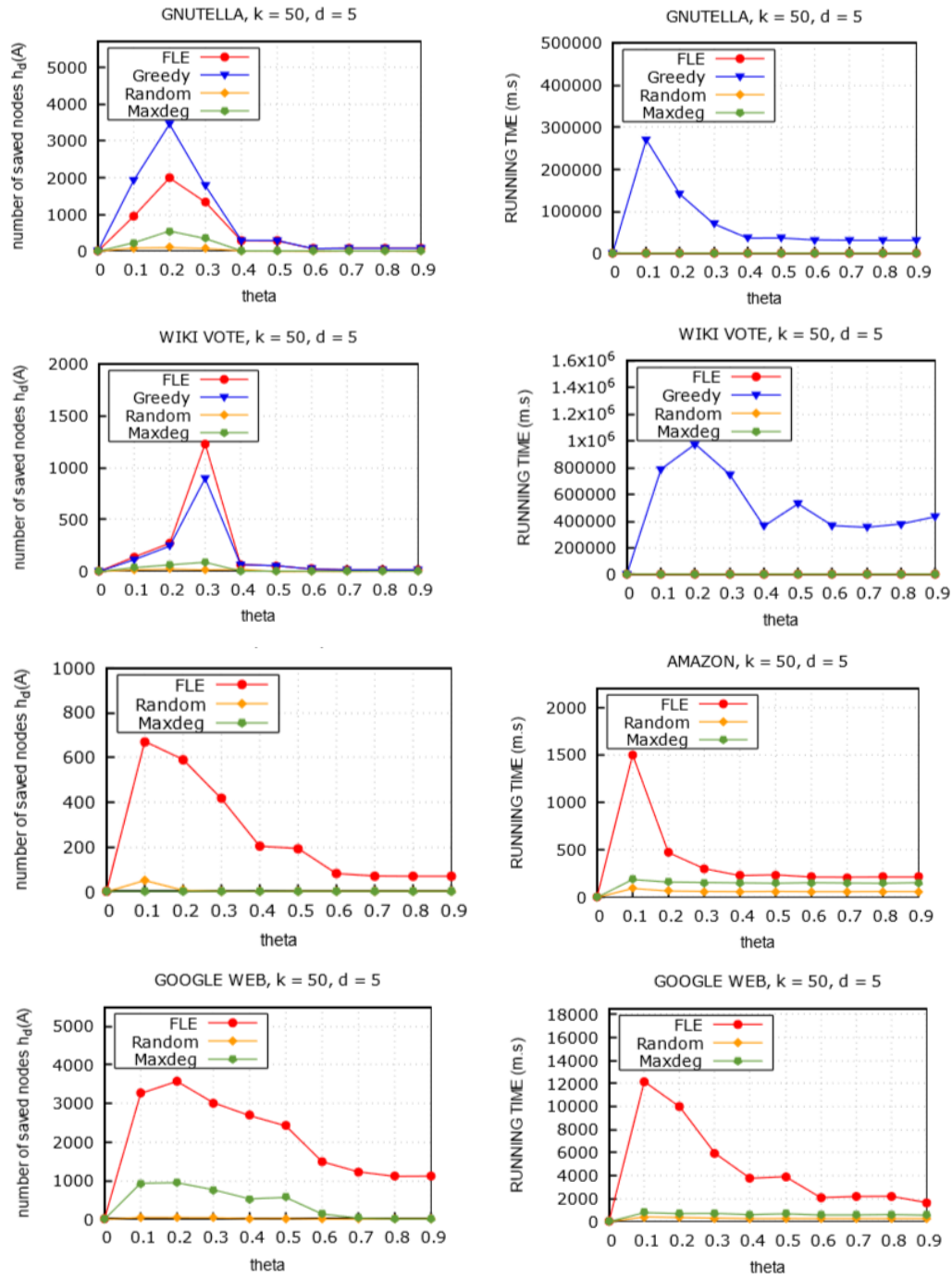
Chúng tôi cũng xem xét ảnh hưởng của tham số  $\theta$  đối với các thuật toán bằng cách



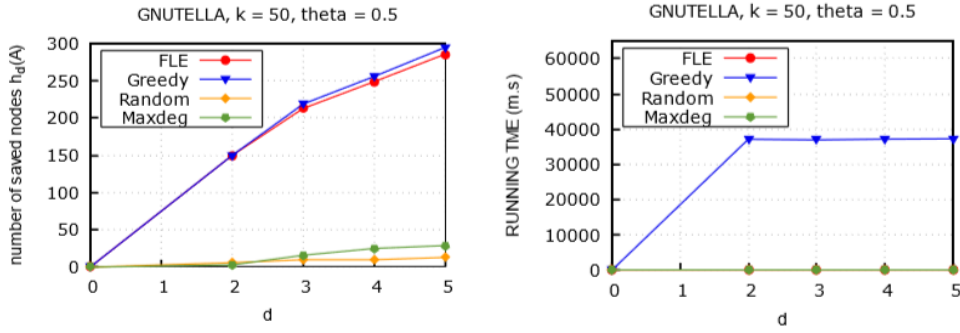
cho  $\theta$  thay đổi trong khi giữ nguyên  $d$  và  $k$ . Chúng tôi lựa chọn giá trị  $d = 5$ ,  $k = 50$ . Đối với mạng Amazon và Google Web, số đỉnh cứu được sẽ giảm nếu giá trị  $\theta$  giảm. Đối với mạng Gnutella và Wiki Vote, số đỉnh cứu được tăng khi  $\theta$  tăng từ 0.1 đến 0.3 và giảm khi  $\theta$  tăng từ 0.3 đến 0.9. Tổng quát lại, ta có thể thấy rằng thực tế việc giá trị  $\theta$  càng cao sẽ làm quá trình lây lan gặp khó khăn. Từ hình 3.3, ta thấy Greedy và FLE cho kết quả tốt hơn nhiều so với hai thuật toán còn lại. Điều này một lần nữa cho thấy tính ưu việt của thuật toán FLE.



Hình 3.2: So sánh chất lượng lời giải và thời gian chạy của các thuật toán khi  $k$  thay đổi,  $d = 5$ ,  $\theta = 0.5$ .



Hình 3.3: So sánh chất lượng lời giải và thời gian chạy của các thuật toán khi  $\theta$  thay đổi,  $d = 5, k = 50$ .



Hình 3.4: So sánh chất lượng lời giải và thời gian chạy của các thuật toán khi  $d$  thay đổi,  $\theta = 0.5$ ,  $k = 50$  trong mạng Gnutella.

## 2. Bài toán 2: Tìm chi phí nhỏ nhất để ngăn chặn thông tin sai lệch

Với cách tiếp cận như cách Twitter đã làm năm 2016 là loại bỏ 125.000 tài khoản được nghi ngờ có liên quan đến các hoạt động khủng bố, hay cách tiếp cận như Facebook đã xóa 30.000 tài khoản giả mạo lan truyền các tin đồn trước cuộc bầu cử tổng thống Pháp năm 2017, bài toán Targeted Misinformation Blocking được đưa ra nghiên cứu nhằm mục đích tìm giải pháp xác định được những tài khoản cần loại bỏ khỏi mạng để hạn chế sự phát tán thông tin sai lệch một cách tối ưu nhất.

### 2.1. Tổng quan bài toán

Nguồn phát tán thông tin sai lệch có thể được phát hiện qua các phương pháp khảo sát, khai phá dữ liệu [24, 25] và được biểu diễn dưới dạng là tập hợp các đỉnh. Theo các nghiên cứu mới đây [41, 43, 44], chiến thuật hiệu quả trong việc ngăn chặn phát tán thông tin sai lệch là xóa bỏ đỉnh hoặc cạnh có vai trò quan trọng trong việc lan truyền thông tin sai lệch. Song tính khả thi của các nghiên cứu trước đây lại bị hạn chế bởi một số khía cạnh sau. Thứ nhất họ tập trung vào việc ngăn ngừa thông tin sai lệch bằng nguồn ngân sách hạn chế. Điều này rõ ràng là không thực tế bởi chúng ta khó có thể xác định được bao nhiêu ngân sách cho đủ để ngăn chặn sự lây lan sao cho hiệu quả. Thứ hai, để đảm bảo độ tin cậy của mạng xã hội, chúng ta cần có một biện pháp "khử nhiễm" thông tin sai lệch với một ngưỡng cho trước, như tỉ lệ giữa người dùng bị nhiễm và người dùng được khử nhiễm. Do đó bài toán này được đưa ra với tính thực tế cao hơn đó là: dùng ít ngân sách nhất sao cho số lượng người dùng được khử nhiễm lớn hơn một ngưỡng cho trước. Giải pháp đưa ra nhằm tìm một tập đỉnh có ít phần tử nhất để loại bỏ khỏi đồ thị để sao cho nguồn phát tán thông tin sai lệch không thể lan truyền tới được ít nhất ngưỡng  $\gamma$  đỉnh yêu cầu. Tổng quan về giải pháp của chúng tôi được tóm lược như sau:

- Đầu tiên chúng tôi định nghĩa bài toán Xác định và ngăn chặn thông tin sai lệch (Targeted Misinformation Blocking - TMB). Và sau đó chứng minh bài toán TMB

thuộc lớp bài toán  $\#P$  – khó.

- Với hàm mục tiêu chúng tôi chứng minh nó là hàm submodular và monotone. Trên cơ sở đó, đề xuất một thuật toán tham lam với tỉ lệ  $1 + \ln(\gamma/\epsilon)$ . Song do những hạn chế của thuật toán tham lam, chúng tôi xây dựng thuật toán Scalable TMB for LT model Algorithm (STMB) để giải quyết cho bài toán TMB với những dữ liệu lớn hơn.
- Việc thực nghiệm được triển khai trên các bộ dữ liệu đã công bố thu thập từ nguồn uy tín và một bộ dữ liệu được chúng tôi thu thập từ mạng xã hội Facebook.

## 2.2. Mô hình lan truyền thông tin Ngưỡng tuyến tính Linear Threshold

Bài toán này sử dụng mô hình lan truyền thông tin Ngưỡng tuyến tính Linear Threshold. Mô hình lan truyền này là mô hình phổ biến, thường được sử dụng trong các nghiên cứu về mạng xã hội. Chi tiết về mô hình LT đã được trình bày trong chương 2, ta có thể tóm gọn quá trình lan truyền thông tin trên mô hình LT như sau:

Trong mô hình LT, mỗi đỉnh  $v \in V$ , có hai trạng thái active và inactive. Quá trình lan truyền trên mô hình LT được diễn tả như sau: Đầu tiên, mỗi đỉnh  $v$  được gán một trọng số ngẫu nhiên  $\theta_v \in [0, 1]$  thể hiện giá trị active của đỉnh  $v$  phải đạt được để có thể active  $v$ . Tiếp đến quá trình lan truyền diễn ra trong các vòng  $t = 1, 2, 3, \dots$  như sau:

- Ở vòng 1, tất cả các đỉnh  $s_i \in S$  ở trạng thái active và các đỉnh còn lại ở trạng thái inactive.
- Ở vòng  $t > 1$ , một đỉnh  $v$  ở trạng thái inactive được active nếu thỏa mãn điều kiện

$$\sum_{u \in N_{in}(v) \text{ và } u \text{ là active}} w(u, v) \geq \theta_v$$

- Khi một đỉnh ở trạng thái active thì sẽ luôn duy trì trạng thái đó trong suốt quá trình lan truyền thông tin. Quá trình lan truyền kết thúc khi không có đỉnh nào được active thêm.

## 2.3. Định nghĩa bài toán

Kí hiệu  $G = (V, E, w)$  là một đồ thị có hướng biểu diễn cho một mạng xã hội trên mô hình LT với  $V$  biểu diễn tập người dùng,  $E$  biểu diễn mối quan hệ giữa các cặp người người dùng,  $w$  là tập trọng số của các cạnh biểu diễn tần số tương tác giữa các người dùng. Gọi  $N_{out}(v)$  và  $N_{in}(v)$  lần lượt là tập đỉnh đi ra từ  $v$  và đi vào  $v$ . Mỗi cạnh có hướng  $(u, v) \in E$  được gán một trọng số  $w(u, v) \in [0, 1]$  đảm bảo điều kiện  $\sum_{u \in N_{in}(v)} w(u, v) \leq 1$ .

Đưa ra một tập nguồn  $S = s_1, s_2, \dots, s_n, s_i \in V$  là tập tiềm năng phát tán thông tin sai lệch (tương tự như tập hạt giống trong bài IM [16]).

Kí hiệu  $\sigma_S(G)$  là ảnh hưởng lan truyền của tập  $S$  trên đồ thị  $G$  dưới mô hình LT hay là số lượng đỉnh được active bởi nguồn  $S$ . Theo nghiên cứu của Kempe và các cộng sự [15] đã chỉ ra rằng mô hình LT là tương đương với khả năng tiếp cận trong một đồ thị ngẫu nhiên  $g$ , thường được gọi là đồ thị live-edge hoặc đồ thị mẫu. Trên mô hình LT, đồ thị live-edge được xây dựng như sau:

- Với mỗi  $v \in V$ , chọn tối đa một cạnh đi vào nó một cách ngẫu nhiên, xác suất cạnh  $(u, v)$  được chọn là  $w(u, v)$ .
- Xác suất không có cạnh nào được chọn là  $1 - \sum_{u \in N_{in}(v)} w(u, v)$ .

Cạnh được chọn gọi là cạnh sống (live-edge) và tất cả các cạnh khác gọi là cạnh bị chặn (blocked-edge). Chúng ta có xác suất tạo ra một đồ thị live-edge  $g$  từ  $G$  là:

$$\Pr[g|G] = \prod_{v \in V} p(v, g, G) \quad (3.1)$$

với

$$p(v, g, G) = \begin{cases} w(u, v), & \text{Nếu } \exists u : (u, v) \in E_g \\ 1 - \sum_{u: (u, v) \in E} w(u, v), & \text{Trường hợp khác} \end{cases} \quad (3.2)$$

Ảnh hưởng lan truyền của nguồn  $S$  khi xóa bỏ các đỉnh trong tập  $A$  trên đồ thị  $G[V \setminus A]$  kí hiệu là  $\sigma_S(G \setminus A)$ . Với  $A = \emptyset$ , trong [15] đã chỉ ra ảnh hưởng lan truyền của tập nguồn  $S$  trên  $G$  được tính bởi công thức:

$$\sigma_S(G) = \sum_{g \in X_G} \Pr[g|G] |R(g, S)| \quad (3.3)$$

trong đó  $R(g, S)$  là tập các đỉnh đến được từ  $S$  trên  $g$  và  $X_G$  là tập tất cả các đồ thị live-edge được tạo ra từ  $G$ . Đặc biệt, ảnh hưởng lan truyền của tập  $S$  có thể được tính bằng đường đi đơn bắt đầu từ  $u \in S$  theo công thức:

$$\sigma_S(G) = \sum_{x \in P(G, S)} \prod_{e \in x} w(e) \quad (3.4)$$

với  $P(G, S)$  là tập tất cả các đường đi đơn từ  $s \in S$  tới bất kỳ đỉnh  $v \in G \setminus \{S \cup \{s\}\}$  trên đồ thị  $G$ .

**Phát biểu bài toán (Bài toán TMB):** Gọi  $G(V, E, w)$  là một đồ thị có hướng biểu diễn một mạng xã hội. Đưa ra một tập nguồn  $S = s_1, s_2, \dots, s_k, S \in V$  là tập tiềm năng phát tán thông tin sai lệch và một số nguyên  $\gamma < |V|$  thể hiện cho số lượng người dùng tối thiểu chúng ta muốn đảm bảo thông tin sai lệch không lan truyền đến được họ.

**Yêu cầu:** Tìm tập đỉnh  $A \in V \setminus S$  có ít đỉnh nhất sao cho khi xóa các đỉnh trong tập  $A$  khỏi đồ thị  $G$  thì đảm bảo  $h(A) = \sigma_S(G) - \sigma_S(G \setminus A) \geq \gamma$ , tức là đảm bảo ít nhất  $\gamma$  người dùng không tiếp xúc với thông tin sai lệch.

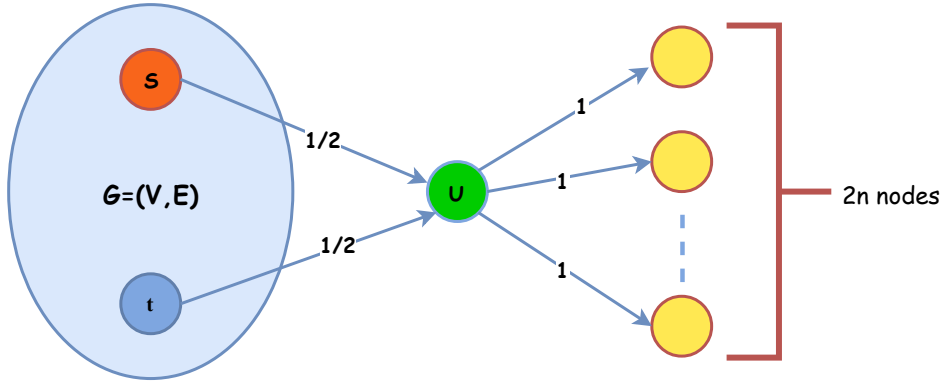
## 2.4. Độ phức tạp của bài toán

Trong mục này chúng tôi chỉ ra một số kết quả về độ khó của bài toán TMB trên mô hình LT.

**Định lý 3.2.** Cho nguồn tiềm năng phát tán thông tin sai lệch  $S$  dưới mô hình LT, bài toán TMB là  $\#P$  – khó nếu nguồn  $S$  chỉ chứa 1 đỉnh.

*Chứng minh.* Để chứng minh TMB là  $\#P$  – khó chúng tôi sẽ chỉ ra độ khó của bài toán TMB là lớn hơn hoặc bằng với độ khó của bài toán  $s$ - $t$  paths mà đã được chứng minh trong [37] là có độ khó  $\#P$  – khó.

**Định nghĩa 3.2** (Bài toán  $s$ - $t$  paths [37]). Cho một đồ thị có hướng  $G = (V, E)$ ,  $|V| = n$ ,  $|E| = m$ , bài toán  $s$ - $t$  paths yêu cầu tìm số đường đi có hướng từ đỉnh  $s$  tới đỉnh  $t$  mà đi qua mọi đỉnh nhiều nhất một lần.



Hình 3.5: Giảm thể từ bài toán  $s$ - $t$  paths sang bài toán TMB.

Xem xét thể hiện  $\mathcal{I}_1$  của bài toán  $s$ - $t$  paths, với  $G = (V, E)$ ,  $s, t$  cho trước. Như minh họa trong hình 3.5, từ  $G$ , chúng ta xây dựng  $G'$  như sau: Thêm đỉnh  $u$  và hai cạnh  $(s, u)$ ,  $(t, u)$  với trọng số  $w(s, u) = w(t, u) = 1/2$ , tiếp đó thêm tập  $Q$  gồm  $2n$  đỉnh và kết nối  $u$  với chúng với trọng số là 1. Với các cạnh khác, thiết lập trọng số  $w = 1/\Delta$ , trong đó  $\Delta$  là bậc vào lớn nhất của đồ thị  $G$ . Chúng ta chia tất cả các đường đi bắt đầu từ  $s$  thành 2 nhóm: Nhóm thứ nhất là tập tất cả các đường đi có đỉnh kết thúc thuộc tập  $Q$  và nhóm

thứ hai là tập còn lại. Từ công thức (3.4), chúng ta có:

$$\begin{aligned}
\sigma_S(G') &= \sum_{x \in P(G', s)} \prod_{e \in x} w(e) = \sum_{v \in G' \setminus Q} \left( \sum_{x \in P(G', s, v)} \prod_{e \in x} w(e) \right) + \sum_{v \in Q} \left( \sum_{x \in P(G', s, v)} \prod_{e \in x} w(e) \right) \\
&= \sum_{x \in P(G, s)} \prod_{e \in x} w(e) + \sum_{x \in P(G', s, u)} \prod_{e \in x} w(e) + \sum_{v \in Q} \left( \sum_{x \in P(G', s, v)} \prod_{e \in x} w(e) \right) \\
&= \sigma_S(G) + \sum_{x \in P(G', s, u)} \prod_{e \in x} w(e) + \sum_{v \in Q} \left( \sum_{x \in P(G', s, v)} \prod_{e \in x} w(e) \right) \tag{3.5}
\end{aligned}$$

và,

$$\begin{aligned}
\sum_{v \in Q} \left( \sum_{x \in P(G', s, v)} \prod_{e \in x} w(e) \right) &= \sum_{v \in Q} \left( \sum_{x \in P(G', s, u)} w(u, v) \prod_{e \in x} w(e) \right) \\
&= \sum_{v \in Q} \left( \sum_{x \in P(G', s, u)} \prod_{e \in x} w(e) \right) \quad (\forall i \ w(u, v) = 1) \\
&= 2n \cdot \sum_{x \in P(G', s, u)} \prod_{e \in x} w(e) \tag{3.6}
\end{aligned}$$

Từ (3.5) và (3.6), chúng ta thu được:

$$h(u) = \sigma_S(G') - \sigma_S(G' \setminus \{u\}) = \sigma_S(G') - \sigma_S(G) = (2n + 1) \cdot \sum_{x \in P(G', s, u)} \prod_{e \in x} w(e) \tag{3.7}$$

$$= (2n + 1) \cdot \left( w(s, u) + \sum_{x \in P(G, s, t)} w(t, u) \prod_{e \in x} w(e) \right) \tag{3.8}$$

$$= (2n + 1) \cdot \left( \frac{1}{2} + \frac{1}{2} \sum_{x \in P(G, s, t)} \prod_{e \in x} w(e) \right) \quad (\forall i \ w(s, u) = w(t, u) = \frac{1}{2}) \tag{3.9}$$

$$= \frac{2n + 1}{2} \left( \sum_{x \in P(G, s, t)} \prod_{e \in x} w(e) + 1 \right) \tag{3.10}$$

$$= \frac{2n + 1}{2} \left( \sum_{i=0}^{n-1} \alpha_i w^i + 1 \right) \tag{3.11}$$

với  $\alpha_i = |P_i(G, s, t)|$ . Đặt  $f(w) = \sum_{i=0}^{n-1} \alpha_i w^i$ . Từ cấu trúc liên kết của  $G'$ , chúng ta thấy rằng  $h(u) = \max_{v \in G'} h(v)$  và  $0 \leq f(w) < 1$ . Nếu có thể xác định  $f(u) \geq \beta$  với  $\forall \beta \in [0, 1]$  trong thời gian đa thức thì chúng ta sẽ giải được bài toán  $s$ - $t$  paths trong thời gian đa thức. Sử dụng tìm kiếm nhị phân trên đoạn  $[1, \Delta^{n-1}]$  chúng ta có thể xác định được giá trị của  $f(w)$ . Yêu cầu này được thực hiện trong  $\mathcal{O}(\log(\Delta^{n-1})) = \mathcal{O}((n-1) \log \Delta) = \mathcal{O}(n \log n)$ . Do đó việc tính toán  $f(u)$  có thể thực hiện trong thời gian đa thức. Tiếp đến thay đổi trọng số  $w$  thành  $n$  giá trị phân biệt  $\frac{1}{\Delta}, \frac{1}{\Delta+1}, \dots, \frac{1}{\Delta+n-1}$ , với mỗi giá trị  $w$  áp dụng phương pháp

trên chúng ta thu được giá trị  $f(w)$  tương ứng. Từ đó có được  $n$  phương trình  $\sum_{i=0}^{n-1} \alpha_i w^i = f(w)$ ,  $w \in \{\frac{1}{\Delta}, \frac{1}{\Delta+1}, \dots, \frac{1}{\Delta+n-1}\}$  với  $\{\alpha_0, \alpha_2, \dots, \alpha_{n-1}\}$  là các biến số. Ma trận tương ứng với hệ  $n$  phương trình trên là  $M_{n \times n} = \{m_{ij}\}$  and  $m_{ij} = w^i$ ,  $i, j = 0, \dots, n-1$ , dễ thấy đây là ma trận Vandermonde do đó có thể dễ dàng tìm được lời giải  $\{\alpha_0, \alpha_2, \dots, \alpha_{n-1}\}$ . Tổng số  $s$ - $t$  paths trên đồ thị  $G$  là  $\sum_{i=0}^{n-1} \alpha_i$ . Vì thế, nếu chúng ta xác định được  $f(u) \geq \beta$  với bất kì  $\beta \in [0, 1]$  trong thời gian đa thức thì sẽ giải được bài toán  $s$ - $t$  paths trong thời gian đa thức.

Bây giờ, xét thể hiện  $\mathcal{I}_2$  của bài toán TMB trong đó  $S = \{s\}$ ,  $\gamma = \frac{2n+1}{2}(\beta + 1)$ ,  $\beta \in [0, 1]$ . Giả sử rằng  $\mathcal{A}$  là một thuật toán giải quyết bài toán TMB trong thời gian đa thức. Xem xét 2 trường hợp: (1) Nếu  $\mathcal{A}$  trả ra kết quả là tập  $A$  có  $|A| = 1$ , thì chúng ta chỉ cần chọn  $A = u$ . Kết luận  $f(w) \geq \beta$ ; (2) Nếu  $\mathcal{A}$  trả ra kết quả tập  $A$  có  $|A| \geq 1$  thì ngoài  $u$  sẽ cần chọn thêm một số đỉnh khác. Kết luận  $f(w) < \beta$ . Vì vậy  $\mathcal{A}$  có thể được sử dụng để so sánh  $f(w)$  với  $\beta$  và có thể giải quyết bài toán  $s$ - $t$  paths. Điều này chỉ ra rằng bài toán TMB là khó hơn hoặc bằng bài toán  $s$ - $t$  paths.  $\square$

## 2.5. Thuật toán đề xuất

### a) Thuật toán tham lam (Greedy algorithm)

Trong phần này chúng tôi giới thiệu một thuật toán xấp xỉ đảm bảo tỉ lệ  $1 + \ln(\gamma/\epsilon)$  dựa vào tính chất submodular và tính monotone của hàm  $h(\cdot)$ . Hàm  $f(x)$  được gọi là hàm submodular nếu  $\forall A \subset T, v \notin T$   $f(A + \{v\}) - f(A) \geq f(T + \{v\}) - f(T)$ . Còn  $f(x)$  được gọi là monotone nếu  $\forall T \subseteq S$  chúng ta có  $f(T) \leq f(S)$ .

---

#### Algorithm 4: Greedy Algorithm (GA)

---

**Data:** Graph  $G = (V, E, w)$ ,  $S = \{s_1, s_2, \dots, s_k\}$ , threshold  $\gamma < |V|$ , parameter

$\epsilon \in (0, \gamma)$

**Result:** set of nodes  $A$

```

1  $A \leftarrow \emptyset;$ 
2 while  $h(A) < \gamma - \epsilon$  do
3    $u = \arg \max_{v \in V \setminus \{A \cup S\}} h(A + \{v\}) - h(A);$ 
4    $A \leftarrow A \cup \{u\};$ 
5 end
6 return  $A;$ 
```

---

**Định lý 3.3.** Hàm  $h(\cdot)$  là hàm submodular và monotone.

*Chứng minh.* Theo định lý 5 trong [41], cho  $A \subseteq T$  chúng ta có  $h(T) - h(A) = \sigma_S(G) - \sigma_S(G \setminus T) - (\sigma_S(G) - \sigma_S(G \setminus A)) = \sigma_S(G \setminus A) - \sigma_S(G \setminus T) \geq 0$ . Do đó  $h(\cdot)$  là một hàm monotone.



Tiếp đến chúng ta chỉ ra rằng  $\sigma_S(G \setminus A)$  là một hàm submodular với  $A$  là biến số. Kí hiệu  $N_E(A)$  là tập tất cả các cạnh  $(u, v) \in E$  với  $u, v \in A$ . Đặt  $E_{T,v} = N_E(T + \{v\}) \setminus N_E(T)$ ,  $E_{A,v} = N_E(A + \{v\}) \setminus N_E(A)$  để thấy  $E_{T,v} \subseteq E_{A,v}$  nếu  $A \subseteq T \Rightarrow N_E(A) \cup E_{T,v} \subseteq N_E(A + \{v\})$ . Lấy  $\sigma_S(G \setminus N_E(X))$  là ảnh hưởng lan truyền của tập  $S$  trên đồ thị  $G$  sau khi loại bỏ tập cạnh  $N_E(X) \subset E$ , chúng ta nhận thấy  $\sigma_S(G \setminus X) = \sigma_S(G \setminus N_E(X))$ . Theo định lý 6 trong [41],  $\forall N_E(X) \subseteq N_E(Y), e \in N_E(Y) \setminus N_E(X)$ , chúng ta có:

$$\sigma_S(G \setminus (N_E(X) \cup \{e\})) - \sigma_S(G \setminus X) \leq \sigma_S(G \setminus (N_E(Y) \cup \{e\})) - \sigma_S(G \setminus Y) \quad (3.12)$$

Vì thế,

$$\sigma_S(G \setminus A) - \sigma_S(G \setminus (A \cup \{v\})) = \sigma_S(G \setminus N_E(A)) - \sigma_S(G \setminus N_E(A + \{v\})) \quad (3.13)$$

$$\geq \sigma_S(G \setminus N_E(A)) - \sigma_S(G \setminus (N_E(A) \cup E_{T,v})) \quad (3.14)$$

$$\geq \sigma_S(G \setminus N_E(T)) - \sigma_S(G \setminus (N_E(T) \cup E_{T,v})) \quad (3.15)$$

$$= \sigma_S(G \setminus T) - \sigma_S(G \setminus (T \cup \{u\})) \text{ (Theo BDT (3.12))} \quad (3.16)$$

Kết hợp với công thức  $h(A) = \sigma_S(G) - \sigma_S(G \setminus A)$  chúng ta có được  $h(\cdot)$  là một hàm submodular.  $\square$

**Định lý 3.4.** Với  $\forall \epsilon \in (0, \gamma)$ , thuật toán 4 đưa ra lời giải  $A$  thỏa mãn  $h(A) \geq \gamma - \epsilon$ , và  $|A|$  là không vượt quá tích của  $1 + \ln(\gamma/\epsilon)$  với số đỉnh của lời giải tối ưu

*Chứng minh.* Chi tiết cách chứng minh định lý 3.4 được chỉ ra trong [22].  $\square$

Thuật toán tham lam có cách tiếp cận khá đơn giản là mỗi bước tìm đỉnh  $u$  sao cho hàm  $h(A)$  tăng nhiều nhất. Song vấn đề khó khăn nhất gặp phải là tính giá trị hàm  $\sigma(\cdot)$  có độ khó là  $\#P$  – khó, chứng minh trong [37]. Vì vậy chúng tôi đưa ra một thuật toán hiệu quả trong phần sau.

### ***b) Thuật toán STMB (STMB algorithm)***

Thuật toán tham lam đưa ra một lời giải tốt xong lại không đảm bảo yêu cầu về thời gian với dữ liệu lớn do đó trong phần này chúng tôi đưa ra một giải pháp đảm bảo yêu cầu cả về chất lượng lời giải cũng như thời gian tính toán. Dưới đây là các bước chính trong thuật toán chúng tôi đề xuất:

- Đầu tiên sắp nhập các đỉnh trong tập nguồn  $S$  thành một nút  $I$  bằng cách áp dụng thuật toán 5 được đề xuất bởi Zhang [45]. Chúng ta thu được một đồ thị mới và nguồn phát tán thông tin sai lệch bây giờ chỉ còn là một đỉnh  $I$  duy nhất.

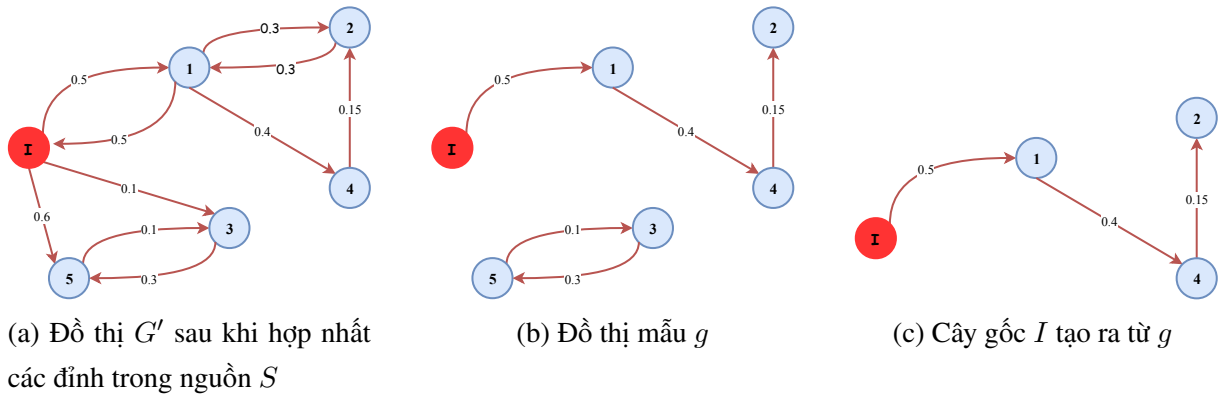
**Algorithm 5:** MERGE Algorithm(MERGE)**Data:** Graph  $G = (V, E, w)$ ,  $S = \{s_1, s_2, \dots, s_k\}$ **Result:** Graph  $G' = (V', E', w')$ ,  $I$ 

```

1  $G' \leftarrow G$ ;
2 Add node  $I$  to  $G'$ 
3 foreach  $s \in S$  do
4   if  $\exists (s, v)$  then
5     if  $(I, v) \notin G'$  then
6       Add edge  $(I, v)$  to  $G'$ 
7        $w'(I, v) = w(s, v)$ 
8     else
9        $w'(I, v) = w'(I, v) + w(s, v)$ 
10    end
11    Remove  $(s, v)$  from  $G'$ 
12  end
13 end
14 Remove all node in  $S$  from  $G'$ 
15 return  $G', I$ ;

```

- Tiếp đến chúng tôi tạo ra  $\mu (\mu \in N^*)$  đồ thị mẫu  $g$  từ  $G'$  bằng cách: mỗi đỉnh  $v \in V$  chọn nhiều nhất một đỉnh kề đi đến nó một cách ngẫu nhiên sao cho xác suất chọn cạnh  $(u, v)$  là  $w(u, v)$  và xác suất không chọn cạnh nào là  $1 - \sum_u w(u, v)$ . Những cạnh được chọn được gọi là cạnh sống (live - edge) và tất cả các cạnh còn lại được gọi là cạnh bị chặn (blocked-edge). Chi tiết cách lấy đồ thị mẫu được mô tả trong thuật toán 6 và trong hình 3.6:



Hình 3.6: Ví dụ tạo ra một cây gốc  $I$  từ  $G$  trên mô hình LT

---

**Algorithm 6:** Get Sample Graph LT Algorithm (GSG\_LT)
 

---

**Data:** Graph  $G = (V, E, w), I$ 
**Result:** Tree  $g$  root at  $I$ 

```

1  $g \leftarrow \emptyset$ 
2 for  $v \in V_G$  do
3    $\gamma_{active} \leftarrow \text{random}(0, 1)$ 
4    $\text{sum\_weight} \leftarrow 0$ 
5   for  $u \in N_{in}(v)$  do
6      $\text{sum\_weight} += w(u, v)$ 
7     if  $\text{sum\_weight} \geq \gamma_{active}$  then
8       Add edge  $(u, v)$  to  $g$ 
9       Break
10    end
11  end
12 end
13  $g \leftarrow \text{DFS}_{tree}(I, g)$ 
14 return  $g$ 

```

---

- Trên cây  $T_I^i$ , ảnh hưởng địa phương của nút  $v$  là số nút con của cây gốc  $v$  kể cả  $v$  trên cây  $T_I^i$ . Áp dụng thuật toán DFS để dàng tìm được số nút con của  $v$ . Ảnh hưởng giảm của nguồn  $S$  khi loại đỉnh  $u$  khỏi  $G$  bằng lượng ảnh hưởng giảm trung bình khi loại  $u$  trên tất cả các cây  $T_I^i$ . Kí hiệu  $h(u, T_I^i)$  là lượng ảnh hưởng giảm trên cây  $T_I^i$  sau khi loại bỏ đỉnh  $u$  khỏi  $T_I^i$ . Áp dụng phương pháp Lazy forward của Leskovec [46] để tìm kiếm lời giải.

---

**Algorithm 7:** Scalable TMB for LT model (TMB) Algorithm
 

---

**Data:** Graph  $G = (V, E, w)$ ,  $\mathcal{S} = \{s_1, s_2, \dots, s_q\}$ , threshold  $\gamma > 0$

**Result:** set of nodes  $A$

```

1   $A \leftarrow \emptyset$ ;  $(G', I) \leftarrow \text{Merge}(G, \mathcal{S})$ .
2  Remove all node,  $I$  can't reach in  $G$ .
3  Generate  $\eta$  sample graphs and set  $\eta$  trees  $\mathcal{L} = \{T_I^1, T_I^2, \dots, T_I^{|\eta|}\}$ 
4  For each  $T_I \in \mathcal{L}$ , calculate  $h(u, T_I)$  for all  $u \in T_I$  (by using DFS algorithm).
5  for  $u \in V$  do
6       $u.\delta(u) \leftarrow \frac{1}{\eta} \sum_{T_I \in \mathcal{L}} h(u, T_I)$ ;  $u.cur \leftarrow 1$ 
7      Insert element  $u$  into  $Q$  with  $u.\delta(u)$  as the key
8  end
9   $h_{max} \leftarrow 0$ ;  $iteration \leftarrow 1$ 
10 while  $h_{max} < \gamma - \epsilon$  do
11      $u_{max} \leftarrow \text{dequence } Q$ 
12     if  $u_{max}.cur = iteration$  then
13          $A \leftarrow A \cup \{u_{max}\}$ ;  $iteration \leftarrow iteration + 1$ 
14         foreach  $T_I \in \mathcal{L}_c$  do
15             If  $u_{max} \in T_I$ , remove node  $u_{max}$  and update  $h(v, T_I), \forall v \in T_I$ .
16         end
17          $h_{max} \leftarrow h_{max} + u_{max}.\delta(u_{max})$ 
18     else
19          $u_{max}.\delta(u_{max}) \leftarrow \frac{1}{\eta} (\sum_{T_I \in \mathcal{L}} h(I, T_I) - \sum_{T_I \in \mathcal{L}} h(I, T_I \setminus u_{max}))$ 
20          $u_{max}.cur = iteration$ ; re-insert  $u_{max}$  into  $Q$ 
21     end
22 end
23 return  $A$ ;
```

---

*Độ phức tạp của thuật toán:* Để đánh giá độ phức tạp của thuật toán trước tiên chúng ta đi phân tích độ phức tạp của các thuật toán con trong nó.

- Thuật toán MERGE có độ phức tạp là  $O(k + |N_{in}(S)|)$ .
- Quá trình tạo ra  $\mu$  đồ thị mẫu dòng 3 có độ phức tạp  $O(\mu(m+n))$ .
- Do  $T_I^i$  là một cây nên việc tính toán  $h(u, T_I^i) \forall u \in T_I^i$  mất  $O(\mu n)$
- Phần Lazy Forward có độ phức tạp là  $O(qmn)$  với  $q$  là số vòng lặp từ dòng 10 đến 23.

Vậy nên toàn bộ thuật toán TMB có độ phức tạp là  $O(\mu(m+qn))$ .

## 2.6. Kết quả thực nghiệm với dữ liệu đã công bố

Trong mục này chúng tôi đưa ra kết quả thực nghiệm của thuật toán STMB với 3 bộ dữ liệu đã công bố của các mạng xã hội và so sánh về chất lượng lời giải, tốc độ thực hiện với một số thuật toán cơ bản khác. Thuật toán được thực hiện bằng ngôn ngữ Python 2.7, sử dụng thư viện hỗ trợ NetworkX<sup>1</sup> và chạy trên máy Linux Server với cấu hình 2.30 GHz Intel<sup>®</sup> Xeon<sup>®</sup> CPU E5-2697, bộ nhớ 128G of RAM DDR4.

### a) Dữ liệu

Với mục đích thực nghiệm, chúng tôi sử dụng 3 bộ dữ liệu mạng thực tế thu thập trên trang [snap.stanford.edu](http://snap.stanford.edu)<sup>2</sup>, Một vài thông tin cơ bản về chúng được đưa ra trong bảng 3.2.

Mạng	NetS	AS	NetHEPT
Số đỉnh	1.5K	6.4K	15.2K
Số cạnh	5.4K	12.5K	32.2K
Bậc trung bình	3.8	7.5	4.2
Số đỉnh nguồn	100	300	1000

Bảng 3.2: Các bộ dữ liệu dùng trong thực nghiệm với giải pháp TMB

NetS[38]: Một mạng lưới các nhà khoa học làm việc trên lý thuyết và thử nghiệm mạng, được biên soạn vào tháng 5/2006.

AS[39]: Biểu đồ các bộ định tuyến bao gồm Internet có thể được sắp xếp thành các biểu đồ con được gọi là Autonomous Systems(AS). Mỗi AS trao đổi lưu lượng với một số hàng xóm (peer). Chúng ta có thể xây dựng một mạng lưới các AS liên kết với nhau dựa trên nhật ký Border Gateway Protocol (BGP). Bộ dữ liệu là kết quả từ dự án University of Oregon Route Views.

NetHEPT[15, 40]: Một mạng lưới hợp tác học thuật được trích xuất từ phần "High Energy Physics-Theory" của trang điện tử arXiv<sup>3</sup> trong khoảng thời gian từ năm 1991–2003. Mỗi đỉnh trong mạng đại diện cho một tác giả và mỗi cạnh kết nối 2 đỉnh cho biết 2 tác giả đã đồng tác giả của một bài báo.

### b) Thiết lập tham số

Chúng tôi sử dụng các thiết lập tham số dưới đây cho việc đánh giá thực nghiệm: Theo các nghiên cứu trước [41, 15, 40], chúng tôi thiết lập trọng số các cạnh của đồ thị

<sup>1</sup>Hướng dẫn sử dụng <https://networkx.github.io/>

<sup>2</sup><https://snap.stanford.edu/data/index.html>

<sup>3</sup><https://arxiv.org/>

trên mô hình LT theo công thức  $w(u, v) = 1/N_{in}(v)$ . Với nguồn phát tán thông tin sai lệch chúng tôi chọn ngẫu nhiên từ 4-6% số lượng đỉnh của cả đồ thị, phụ thuộc vào kích thước của đồ thị.

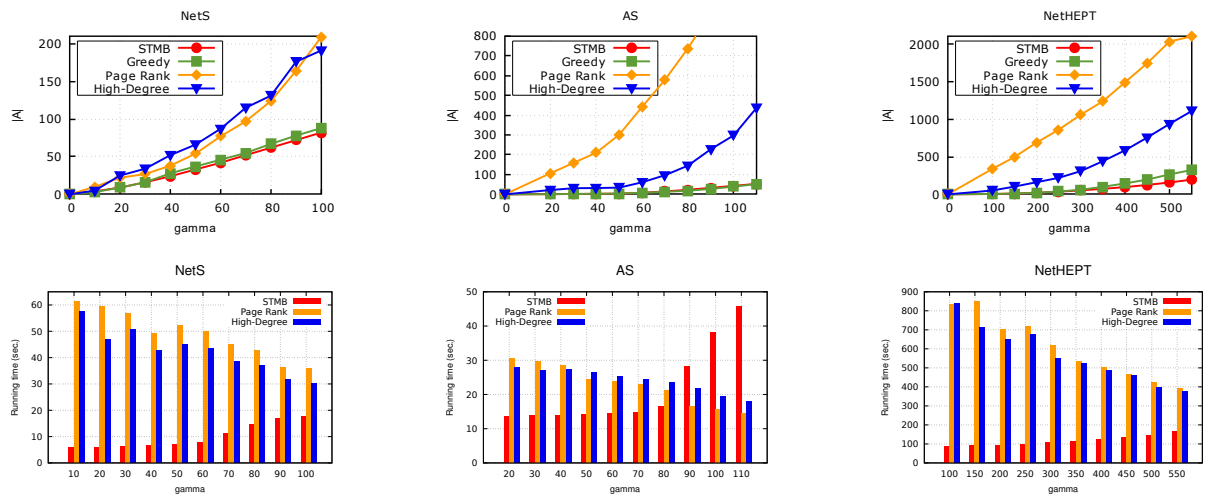
### ***c) Các thuật toán sử dụng***

Trong thực nghiệm, chúng tôi so sánh thuật toán STMB với một số thuật toán sau, các thuật toán này thường được sử dụng trong các nghiên cứu liên quan về vấn đề nhóm tác giả nghiên cứu:

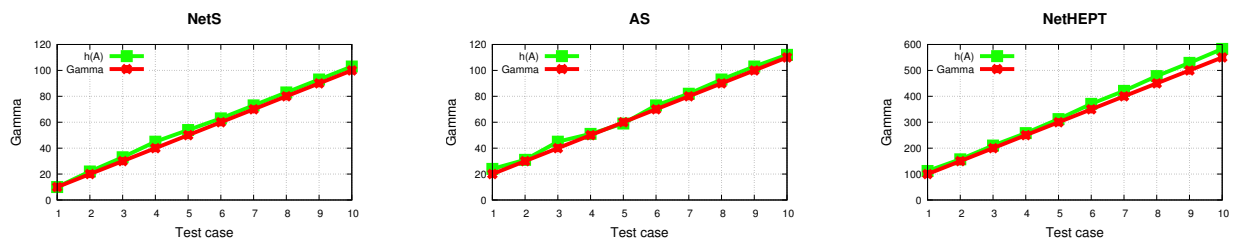
- PageRank: Tính toán ranking của các đỉnh trong đồ thị  $G$  dựa vào cấu trúc các cạnh liên kết với nó. Đây là thuật toán được thiết kế để xếp hạng các trang web từ đó đưa ra kết quả tìm kiếm. Do  $h(\cdot)$  là hàm monotone nên sau khi sắp xếp hạng của các đỉnh trong đồ thị, chúng tôi sử dụng thuật toán tìm kiếm nhị phân để tìm kiếm tập lời giải  $A$  với  $|A|$  đỉnh có rank lớn nhất.
- High-Degree: Một thuật toán heuristic dựa vào bậc của một đỉnh. Bậc của một đỉnh ở đây là tổng số đỉnh được tính theo công thức  $d(v) = N_{in}(v) + N_{out}(v)$ . Tương tự như thuật toán PageRank sau khi sắp xếp các đỉnh dựa trên bậc của chúng, chúng tôi sử dụng thuật toán tìm kiếm nhị phân để tìm lời giải cho bài toán.
- Greedy: Thuật toán Greedy đưa ra trong thuật toán 4 kết hợp với các tối ưu trong [42] giúp tăng tốc độ thuật toán song không làm ảnh hưởng đến chất lượng lời giải.

### ***d) Kết quả thực nghiệm***

Như mô tả trong hình 3.7, số lượng đỉnh được chọn đưa ra bởi thuật toán STMB là nhỏ nhất. STMB tốt hơn lên đến 39% so với thuật toán Greedy, hơn 60%-95% và 57%-87% so với thuật toán PageRank và High-Degree tương ứng. Để kiểm tra kết quả thu được từ thuật toán STMB, chúng tôi chạy 10000 mẫu Monte-Carlo để ước lượng giá trị của hàm  $h(A)$  và kết quả kiểm tra được đưa ra trong hình 3.8. Trong hầu hết các trường hợp  $h(A)$  vượt qua ngưỡng yêu cầu.



Hình 3.7: So sánh chất lượng lời giải và thời gian chạy của các thuật toán cho bài toán TMB



Hình 3.8: Kiểm tra kết quả của thuật toán STMB

	STMB	Greedy	Page Rank	High-Degree
NetS	<b>17.57(s)</b>	14206.80(s)	35.73(s)	30.24(s)
AS	45.70(s)	14074.87(s)	<b>14.39(s)</b>	17.85(s)
NetHEPT	<b>165.12(s)</b>	582566.74(s)	392.34(s)	374.66(s)

Bảng 3.3: So sánh thời gian chạy của các thuật toán với trường hợp  $\gamma$  lớn nhất

Thời gian chạy của các thuật toán trên 3 bộ dữ liệu trong trường hợp có  $\gamma$  là lớn nhất được đưa ra trong bảng 3.3 cho thấy thuật toán Greedy có chất lượng lời giải tương đương với thuật toán STMB song thời gian thực nghiệm trên các bộ dữ liệu lại rất không hiệu quả. Trong khi đó thuật toán STMB có thời gian thực hiện khá là nhanh, không khác biệt nhiều so với các thuật toán cơ bản song vẫn đảm bảo có một lời giải tốt.

## CHƯƠNG 4

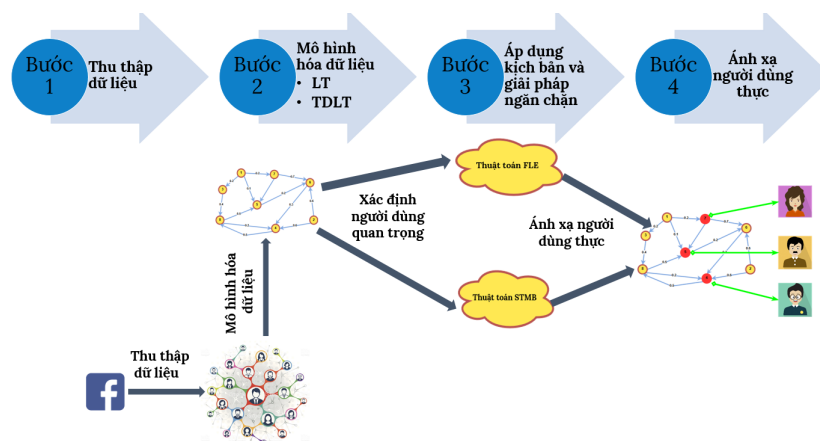
### ÁP DỤNG CÁC GIẢI PHÁP ĐỀ XUẤT TRÊN DỮ LIỆU THỰC

Sau khi đề xuất các kịch bản và các giải pháp lý thuyết ở chương trước, trong nội dung chương này nhóm tác giả tập trung xây dựng ứng dụng cụ thể để giải quyết bài toán hạn chế sự lan truyền của thông tin sai lệch trên mạng xã hội tại Việt Nam.

Nhóm tác giả đề xuất một giải pháp tổng thể bao gồm bốn bước

- Bước đầu tiên: Thu thập dữ liệu. Trong bước này sau khi tiến hành khảo sát nguồn phát tán thông tin sai lệch, nhóm tác giả tiến hành xây dựng công cụ thu thập dữ liệu từ mạng xã hội Facebook xung quanh các nguồn phát tán trên.
- Bước thứ hai: Mô hình hóa dữ liệu. Dựa trên dữ liệu thu được, nhóm tác giả tiến hành mô hình hóa dữ liệu thu được dưới dạng đồ thị, thể hiện quá trình phát tán thông tin bằng hai mô hình tương ứng với hai kịch bản đã đề xuất ở chương trước.
- Bước thứ ba: Áp dụng hai kịch bản ngăn chặn và hai giải pháp tương ứng đã trình bày trong chương 3 đối với dữ liệu đã thu thập và đã được mô hình hóa.
- Bước cuối cùng: Ánh xạ người dùng thực. Bước này tiến hành ánh xạ ngược tập đỉnh thu được trong bước 3 nhằm thu được danh sách người dùng có vai trò quan trọng trong việc lan truyền thông tin sai lệch đối với 2 kịch bản ngăn chặn đã đề xuất.

Hình 4.1 mô tả tổng thể giải pháp nhóm tác giả đề xuất



Hình 4.1: Tổng thể giải pháp



## 1. Thu thập dữ liệu

### 1.1. Xác định nguồn phát tán thông tin sai lệch

Thông tin sai lệch phát tán trên mạng xã hội xuất phát từ các nguồn phát tán với tốc độ nhanh, quy mô rộng lớn. Đã có nhiều nghiên cứu đưa ra nhằm xác định các nguồn phát tán này, tuy nhiên hiệu quả và độ chính xác vẫn còn nhiều hạn chế. Trong thực tế, các nguồn phát tán thông tin sai lệch, các thông tin chống đối và xuyên tạc chính sách, đường lối của Đảng, pháp luật của Nhà nước thường là các đối tượng chủ mưu, cầm đầu, nắm vai trò quan trọng trong các tổ chức phản động. Đây là những đối tượng cơ hội chính trị, đối tượng bất mãn, đối tượng bảo thủ, ngoan cố và rất khó để thuyết phục các đối tượng đó gỡ bỏ các bài đăng, thông tin sai lệch trên mạng xã hội.

Trong nghiên cứu này, nhóm tác giả đã dùng phương pháp khảo sát, điều tra, theo dõi, nghiên cứu tài liệu để tìm ra những đối tượng có tính chất trên. Bằng các biện pháp này nhóm tác giả đã thu thập được nguồn phát tán thông tin sai lệch bao gồm 6 người dùng:

- <https://www.facebook.com/gioan.namphong>: Linh mục giáo xứ Thái Hà Nguyễn Nam Phong, một nhân vật khá nổi tiếng trong làng chống chính quyền trong nước và thường xuyên ra nước ngoài để "liên lạc" với tổ chức khủng bố Việt Tân.
- <https://www.facebook.com/minhnhat.paultran>: Paul Trần Minh Nhật là giáo dân thuộc xứ Ngọc Long, xã Công Thành, Yên Thành, Nghệ An, từng bị bắt và kết án 4 năm tù về tội “hoạt động nhằm lật đổ chính quyền nhân dân”. Hiện tại, Minh Nhật đang là cộng tác viên của trang “Tin mừng cho người nghèo”, mang danh nghĩa rao giảng tin mừng của Chúa, nhưng lại nhuộm màu sắc chính trị, luôn ủng hộ các đối tượng vi phạm pháp luật.
- <https://www.facebook.com/ThaiDung2016>: Gioan Thái Văn Dung là sinh viên tốt nghiệp ngành tin học, đã tham gia quản lý cửa hàng Internet, Do hiểu biết được CNTT mà biết thêm về xã hội, tích cực truyền bá thông tin, hoạt động mạng, tham gia biểu tình chống TQ xâm lược.
- <https://www.facebook.com/pham.doan.trang>: Nhà báo Phạm Đoan Trang, từng công tác tại Báo Pháp Luật thành phố Hồ Chí Minh. Tác giả của cuốn sách “Chính trị Bình Dân” có những nội dung nhạy cảm, mang tính chất kích động, chống phá chính quyền, Đảng và Nhà Nước.
- <https://www.facebook.com/jbnguyenhuuvinh>: Anh Ba Sàm, tên thật là Nguyễn Hữu Vinh là một blogger, từng là công an và đảng viên Đảng Cộng sản Việt Nam,

từng công tác ở Ủy ban Việt kiều Trung ương. Ông bị Chính phủ CHXHCN Việt Nam bắt giữ và bị cáo buộc và phạt tù 5 năm do có hành vi đăng tải các bài viết trên mạng Internet vi phạm Điều 258 Bộ Luật Hình sự năm 2015 sửa đổi bổ sung năm 2017 về Tội lợi dụng các quyền tự do dân chủ xâm phạm lợi ích của Nhà nước, quyền, lợi ích hợp pháp của tổ chức, công dân.

- <https://www.facebook.com/profile.php?id=100015485029386>: Nguyễn Trọng đang sinh sống tại California, Hoa Kỳ. Lợi dụng quyền tự do dân chủ, thường xuyên có những bài viết xuyên tạc, chống phá đường lối chính sách của Đảng, pháp luật của Nhà nước.

## **1.2. Xây dựng công cụ thu thập dữ liệu trên mạng xã hội Facebook**

Dưới những điều kiện khác nhau, như tiềm lực tài chính, khả năng thuyết phục, phạm vi lan truyền thông tin khác nhau, cùng với đó là các mục tiêu khác nhau, có thể kể đến ảnh hưởng lan truyền của thông tin sai lệch, tỉ lệ người dùng nhiễm thông tin sai lệch và với các mô hình lan truyền thông tin khác nhau. Nhóm tác giả đã tiến hành xây dựng công cụ thu thập dữ liệu trực tiếp từ mạng xã hội Facebook, với phạm vi xung quanh các đối tượng chống đối, đối tượng cầm đầu thu được trong bước khảo sát trên.

Facebook cho phép các nhà phát triển (Developers) có thể lấy được thông tin dữ liệu, tuy nhiên điều đó phải được sự cho phép của người dùng thông qua các mã truy cập Access Token (Access Token là một đoạn mã do MXH sinh ra ngẫu nhiên khi người dùng đồng ý cho ứng dụng thứ 3 thực hiện các thao tác đối với tài khoản của người dùng). Điều này là một vấn đề khó khăn khi chúng ta cần khảo sát, lấy thông tin đối với một số lượng lớn người dùng nhằm xây dựng được đồ thị quan hệ giữa các người dùng trên MXH.

Một vấn đề khác là Facebook ngày càng thắt chặt các chính sách bảo mật đối với dữ liệu cá nhân của người sử dụng, trong đó bao gồm cả kiểm soát việc theo dõi thông tin cá nhân của người dùng bởi các người dùng khác. Facebook giới hạn thời gian, quyền truy cập, theo dõi đối với từng tài khoản Facebook, rộng hơn là từng địa chỉ IP trong việc truy cập thông tin cá nhân của người dùng khác.

Nhằm giải quyết vấn đề trên, nhóm tác giả đề xuất một hướng tiếp cận mới là sử dụng công cụ kiểm thử phần mềm tự động mã nguồn mở Selenium cho việc kiểm thử ứng dụng Web. Cụ thể ở đây nhóm tác giả thư viện mã nguồn mở của Selenium trên ngôn ngữ Python, kiểm thử tự động với trình duyệt Chromnium.

### ***a) Tổng quan về công cụ kiểm thử Selenium***

Selenium là một công cụ kiểm thử phần mềm tự động, được phát triển bởi ThoughtWorks từ năm 2004 với tên ban đầu là JavaScriptTestRunner. Đến năm 2007, tác giả Jason Huggins rời ThoughtWorks và gia nhập Selenium Team, một phần của Google và phát triển thành Selenium như hiện nay.

Selenium cung cấp công cụ phát lại (Playback) cho việc tạo ra các bài kiểm thử mà không cần phải học một ngôn ngữ kịch bản khác (Selenium IDE). Nó cũng cung cấp một ngôn ngữ thử nghiệm tên miền cụ thể (Selenese) để tạo ra các bài kiểm thử bằng một số ngôn ngữ lập trình phổ biến bao gồm C#, Groovy, Java, Perl, Python, PHP, Ruby, Scala. Các thử nghiệm sau đó có thể chạy trên hầu hết các trình duyệt Web hiện đại. Selenium triển khai trên các nền tảng Windows, Linux và cả MacOS. Selenium là phần mềm mã nguồn mở, được phát hành theo giấy phép Apache 2.0.

Selenium bao gồm một số thành phần với mỗi thành phần tham gia vào một vai trò cụ thể trong việc hỗ trợ phát triển tự động kiểm tra ứng dụng Web, bao gồm: IDE Selenium, API Client Selenium, Selenium WebDriver, Selenium Remote Control và Selenium Grid. Trong nội dung đề tài nghiên cứu, nhóm tác giả sử dụng hai thành phần chính là API Client Selenium và Selenium WebDriver.

- API Client Selenium Để thay thế cho các bài kiểm tra viết bằng Selen, các bài kiểm tra cũng có thể được viết bằng nhiều ngôn ngữ lập trình khác nhau. Những thử nghiệm này sau đó giao tiếp với Selenium bằng cách gọi các phương thức trong API ứng dụng khách Selenium. Selenium hiện cung cấp các API ứng dụng khách cho Java , C#, Ruby , JavaScript và Python. Với Selenium 2, một API ứng dụng khách mới đã được giới thiệu (với WebDriver là thành phần trung tâm của nó). Tuy nhiên, API cũ (sử dụng lớp Selenium) vẫn được hỗ trợ.
- Selenium WebDriver Selenium WebDriver là sự kế thừa của Selenium RC. Selenium WebDriver chấp nhận các lệnh (được gửi bằng Selenese hoặc thông qua API ứng dụng khách) và gửi chúng tới trình duyệt. Điều này được thực hiện thông qua trình điều khiển trình duyệt cụ thể cho trình duyệt, trình điều khiển sẽ gửi lệnh tới trình duyệt và truy xuất kết quả. Selenium WebDriver không cần một máy chủ đặc biệt để thực hiện các thử nghiệm. Thay vào đó, WebDriver trực tiếp khởi động một cá thể trình duyệt và điều khiển nó. Tuy nhiên, Selenium Grid có thể được sử dụng với WebDriver để thực hiện các kiểm tra trên các hệ thống từ xa. Nếu có thể, WebDriver sử dụng chức năng cấp hệ điều hành gốc hơn là các lệnh JavaScript dựa trên trình duyệt để điều khiển trình duyệt. Điều này bỏ qua các vấn đề với sự khác biệt tinh tế giữa lệnh gốc và JavaScript, bao gồm cả các hạn chế bảo mật.

### ***b) Quá trình thu thập dữ liệu***

Các bước cụ thể của giai đoạn thu thập dữ liệu trên MXH Facebook được mô tả cụ thể như sau:

- Đầu tiên, nhóm tác giả tạo một tài khoản Facebook rồi sử dụng Công cụ kiểm thử tự động Selenium để tự động đăng nhập vào tài khoản Facebook đó.
- Bằng các truy vấn đặc biệt, nhóm tác giả lấy được ID (Identity Number) của tài khoản trên.
- Thông qua phương pháp giám sát và điều tra, thông qua những tin tức, vụ việc trong thời gian gần đây. Nhóm tác giả khảo sát dữ liệu với tập nguồn ban đầu bao gồm 6 tài khoản Facebook đã trình bày ở phần 1.1.
- Tiếp theo, nhóm tác giả tự động truy cập vào trang web có địa chỉ [www.facebook.com/ID/friends](http://www.facebook.com/ID/friends) hoặc [www.facebook.com/name/friends](http://www.facebook.com/name/friends) là trang bạn bè của những người dùng đó.
- Theo thuật toán sắp xếp của Facebook, những người dùng có liên quan nhất và hay tương tác nhất với người dùng đang theo dõi sẽ hiện ra đầu tiên trong trang bạn bè nói trên. Nhóm tác giả sử dụng thư viện BeautifulSoup để lấy mã HTML của trang này, tìm và lấy các đường dẫn chứa địa chỉ Facebook bạn bè của người dùng đang theo dõi, lấy ID của họ.
- Nhóm tác giả thu thập theo phương pháp loang theo chiều rộng: Bắt đầu từ ID Facebook của 6 đối tượng kể trên, từ đó tìm tiếp thông tin về khoảng 20 người bạn xuất hiện đầu tiên trên trang Facebook bạn bè của mỗi người dùng đó. Với mỗi người dùng thu thập được, nhóm tác giả lại tiếp tục tiến hành thu thập như trên, và cứ tiếp tục tìm kiếm đến khi thu thập đủ số lượng người dùng cần thiết.

Cách làm này có ưu điểm là những người dùng thu thập được có nhiều mối liên hệ với nhau, đồ thị MXH thu thập được có nhiều liên kết, mang lại sự trực quan hơn cho cấu trúc đồ thị mạng. Ngoài ra còn có quá trình tiền xử lý, bước này sẽ hạn chế, loại bỏ được những đỉnh cô lập, hay những đỉnh khai thác được ít thông tin (Không công khai danh sách bạn bè).

Bên cạnh những ưu điểm trên, thì cách làm này cũng có nhược điểm là không khai thác được hết thông tin của người dùng, bởi nhiều người dùng không để chế độ công khai bạn bè. Trong thực tế sẽ để lọt nhiều đối tượng nguy hiểm, ẩn chứa nhiều nguy cơ tiềm tàng.

## 2. Mô hình hóa dữ liệu thu được

Ở bước này, từ những dữ liệu thu thập được về danh sách ID người dùng, mối quan hệ bạn bè giữa họ. Nhóm tác giả tiến xử lý, đánh số lại các đỉnh, xây dựng đồ thị mô tả MXH với các đỉnh đại diện cho các người dùng, các cạnh biểu diễn mối quan hệ bạn bè giữa những người đó. Kết quả đầu ra của thuật toán thu thập thông tin trên MXH Facebook và quá trình chuẩn hóa, xử lý bao gồm 2 file dữ liệu:

- ID.txt chứa ánh xạ từ ID của người dùng sang chỉ số đỉnh được đánh số lại. Qua quá trình thu thập dữ liệu, chuẩn hóa, nhóm tác giả thu được 542 ID người dùng, được đánh số lại từ 0 – 541, xem bảng 4.1

0	100003708850657
1	100009945675640
2	100010197936720
3	641613321
4	100002541019308
5	100015485029386
6	100004745583805
...	...
539	100004236103486
540	100004142778267
541	100006259445960

Bảng 4.1: Ánh xạ ID người dùng sang chỉ số được đánh số - ID.txt

- Network.txt chứa mô tả cụ thể của đồ thị MXH đã thu được. File dữ liệu gồm nhiều dòng, mỗi dòng chứa 2 số là chỉ số của 2 người dùng có quan hệ với nhau. Mô tả cụ thể được thể hiện trên bảng 4.2

Từ dữ liệu thu được, nhóm tác giả mô phỏng lại và thu được tổng quan về cấu trúc mạng như trên hình 4.2 :

Tiếp theo, ứng với mỗi kịch bản ngăn chặn, mô hình phát tán thông tin tương ứng, nhóm tác giả tiến hành thiết lập trọng số đại diện cho mức độ tương tác giữa các người dùng với nhau, chuẩn hóa dữ liệu tương ứng với đầu vào yêu cầu của bài toán. Kết quả thu được như sau:

0	7
0	8
0	9
0	4
0	10
0	11
0	12
...	...
99	399
99	495
99	400
99	437
99	498

Bảng 4.2: Mô tả cụ thể của đồ thị MXH thu được - Network.txt



Hình 4.2: Mô phỏng cấu trúc toàn mạng dữ liệu thu được

- Đối với kịch bản FLE: Đầu vào là tệp tin input.inp chứa mô tả của đồ thị mạng xã hội, tập đỉnh nguồn phát tán thông tin sai lệch, số đỉnh cần xóa. File dữ liệu gồm nhiều dòng.
  - + Dòng đầu tiên chứa hai số nguyên dương cách nhau bởi một dấu cách  $nm$ . Trong đó  $n$  là số đỉnh,  $m$  là số cạnh của đồ thị.
  - +  $m$  dòng tiếp theo, mỗi dòng chứa 3 số là chỉ số của 2 đỉnh có cạnh nối với nhau và trọng số cạnh tương ứng.
  - + Dòng tiếp theo là số nguyên  $k$ , số lượng đỉnh nguồn phát tán thông tin, cụ thể  $k = 6$ .
  - +  $k$  dòng tiếp theo, mỗi dòng chứa một số nguyên dương là chỉ số đỉnh là nguồn phát tán thông tin của đồ thị, ở đây các đỉnh nguồn nhóm tác giả đánh số từ 1-6.
  - + Dòng cuối cùng là số nguyên đại diện số lượng đỉnh cần xóa.

Mô tả cụ thể được thể hiện trên bảng 4.3

542	2336	
1	3	0.040
1	5	0.030
...	...	...
542	100	0.062
6		
1		
2		
3		
4		
5		
6		
10		

Bảng 4.3: Dữ liệu đầu vào ứng với giải pháp FLE - input.txt

- Đối với kịch bản STMB: Đầu vào bao gồm 2 tệp tin
  - + Tệp tin TestFacebook.txt chứa mô tả đồ thị mạng xã hội. Tệp tin gồm nhiều dòng, mỗi dòng chứa 3 số là chỉ số của 2 đỉnh có cạnh nối với nhau và trọng số cạnh tương ứng. Mô tả cụ thể được thể hiện trên bảng 4.4

1	3	0.040
1	5	0.030
1	8	0.030
1	9	0.025
...	...	...
539	98	0.062
540	99	0.048
541	100	0.062
542	100	0.062

Bảng 4.4: Dữ liệu đầu vào ứng với giải pháp STMB - TestFacebook.txt

- + Tập tin TestSourceFacebook.txt chứa tập chỉ số các đỉnh nguồn phát tán thông tin. Tập tin gồm 6 dòng, mỗi dòng gồm một số nguyên dương lần lượt từ 1 tới 6, là chỉ số các đỉnh nguồn phát tán thông tin.

### 3. Áp dụng kịch bản và giải pháp ngăn chặn

Trong bước này, nhóm tác giả tiến hành áp dụng hai giải pháp ngăn chặn tương ứng với hai kịch bản đã đề xuất đối với các dữ liệu thực đã thu thập từ mạng xã hội Facebook và đã được mô hình hóa. Kết quả thu được đối với từng giải pháp như sau:

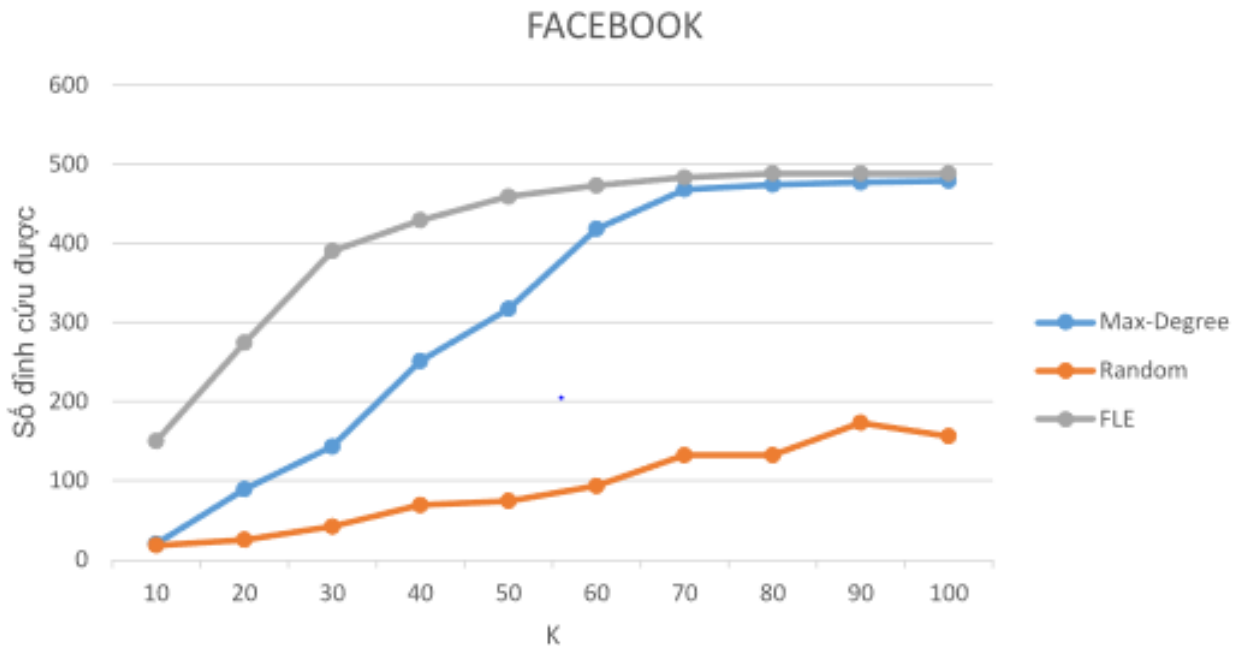
- Áp dụng giải pháp FLE đối với kịch bản LSE. Kết quả được lưu lại trong tập tin Output.txt, được mô tả cụ thể trên bảng 4.5.

8 -> 100002788524244
15 -> 634032152
18 -> 100000315904157
19 -> 1849592060
22 -> 100004460193964
30 -> 100003888766757
39 -> 100000757109103
115 -> 100001730194864
338 -> 100005228346422
397 -> 100003276218628
Saved = 150

Bảng 4.5: Kết quả thực hiện giải pháp FLE với dữ liệu thu thập - Output.txt



So sánh kết quả thực nghiệm này đối với các thuật toán MaxDegree và Random về mặt số lượng đỉnh được cứu, thu được kết quả như hình 4.3:



Hình 4.3: So sánh chất lượng lời giải của các thuật toán với dữ liệu Facebook

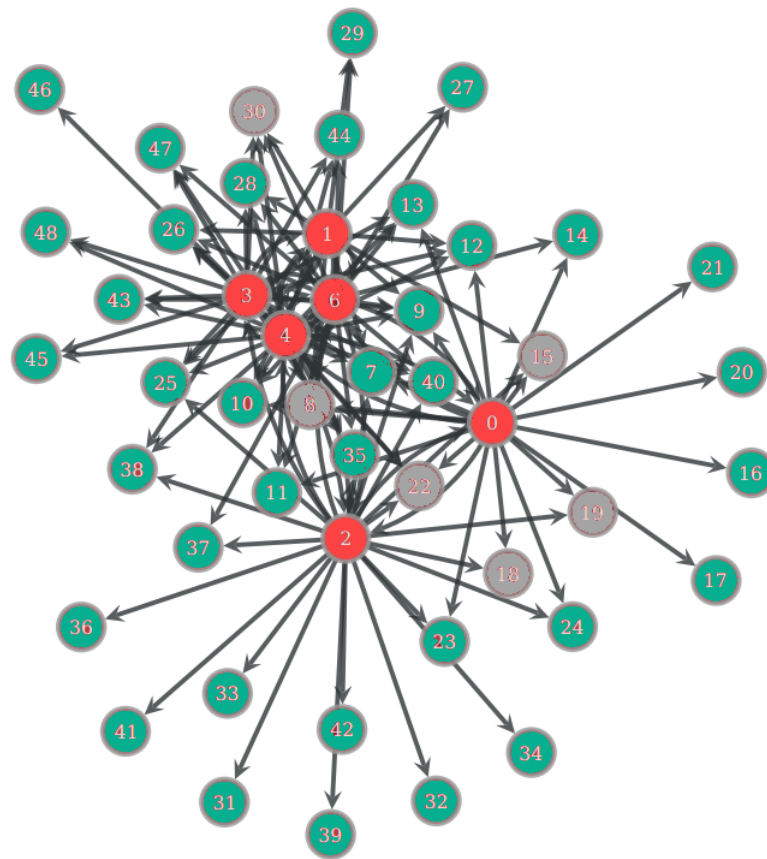
Trong hình 4.3, ta có thể thấy rằng FLE, với cùng một ngưỡng số đỉnh tối đa cứu được chặn là  $k$ , luôn đạt được kết quả tốt hơn so với hai thuật toán còn lại. Đặc biệt, với  $k \geq 80$ , thuật toán FLE chỉ ra được phương pháp chặn bắt để số đỉnh cứu được là tối đa, điều mà hai thuật toán Random và MaxDegree dù với  $k = 100$  cũng không thể làm được.

Vì số đỉnh của đồ thị thu được là khá nhiều cho việc mô tả chi tiết nên chúng tôi lựa chọn một phần của đồ thị để mô phỏng, mà cụ thể là phần trung tâm của đồ thị (phần tập trung nguồn phát tán thông tin sai lệch). Sử dụng công cụ graph\_tool<sup>1</sup>, chúng tôi có thể mô tả lại một phần mạng dữ liệu được sử dụng như hình 4.4:

Các đỉnh được tô màu đỏ thể hiện nguồn phát tán thông tin sai lệch, các đỉnh được tô màu xám là tập đỉnh được lựa chọn để loại ra đồ thị theo kết quả thuật toán, các đỉnh còn lại sẽ được tô màu xanh.

- Áp dụng giải pháp STMB đối với kịch bản TMB. Kết quả được lưu lại trong các tệp tin có cấu trúc tên là Seed\_STMB\_Facebook\_x.y.txt chứa thông tin về các chỉ

<sup>1</sup><https://graph-tool.skewed.de/>



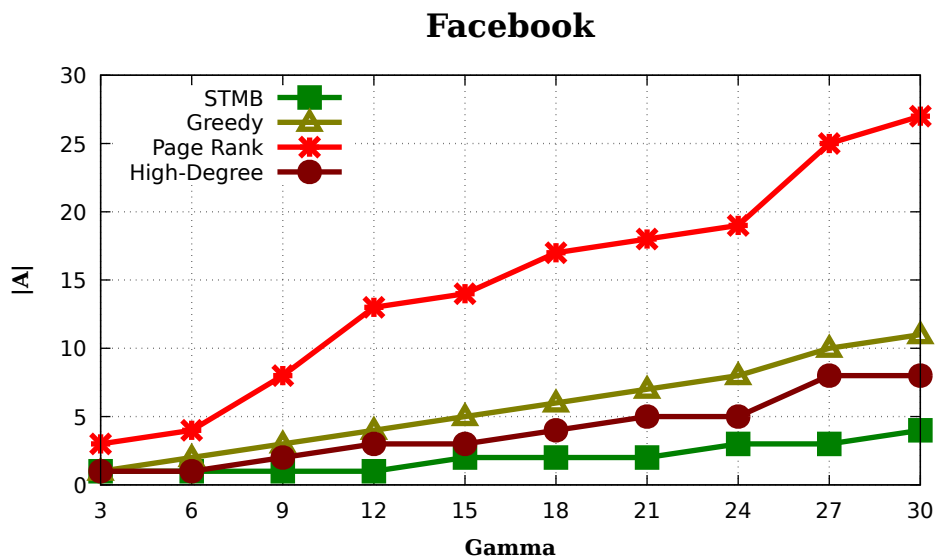
Hình 4.4: Kết quả mô phỏng với dữ liệu thu thập được trên Facebook của thuật toán FLE

số đỉnh tìm được, mỗi đỉnh một dòng, trong đó  $x$  là số lần lấy mẫu,  $y$  là ngưỡng của bài toán đặt ra. Ngoài ra kết quả tổng hợp được lưu trong tệp tin Result.txt, được mô tả cụ thể trên bảng 4.6.

So sánh kết quả thực nghiệm của các thuật toán STMB, Greedy, PageRank, High-Degree về mặt số lượng đỉnh tìm được thỏa mãn yêu cầu (càng ít đỉnh càng tốt), ta thu được hình 4.5

$ V  = 536$
$\sigma_{comming} : 74.84$
$Hmax = 12.841$
$Threshold = 3$
$Hmax = 12.841$
$Threshold = 6$
$Hmax = 12.841$
$Threshold = 9$
...
$Hmax = 28.288$
$Threshold = 24$
$Hmax = 28.288$
$Threshold = 27$
$Hmax = 32.827$
$Threshold = 30$

Bảng 4.6: Kết quả thực hiện giải pháp SMTB với dữ liệu thu thập - Result.txt

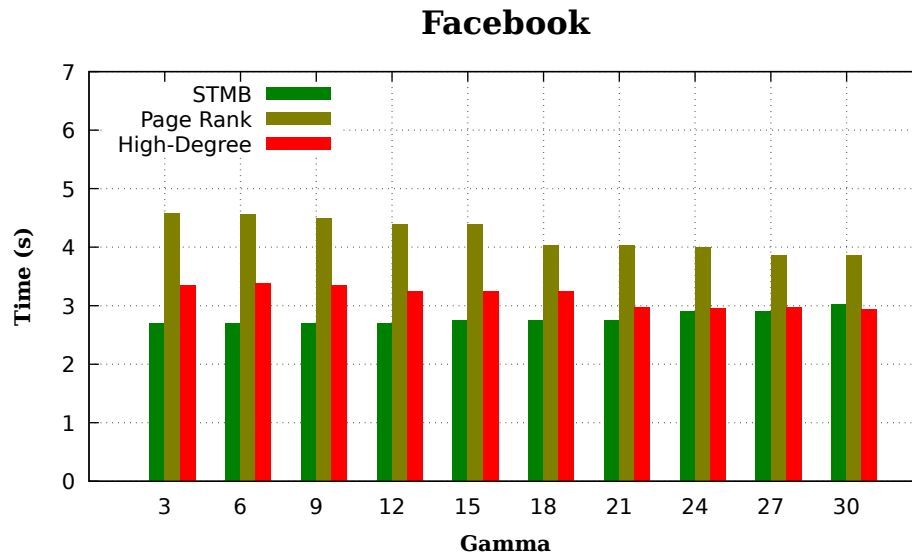


Hình 4.5: So sánh chất lượng lời giải của các thuật toán với dữ liệu thu thập được cho bài toán TMB

Trong hình 4.5, ta có thể thấy rằng số đỉnh tìm được để thỏa mãn yêu cầu của STMB luôn là nhỏ nhất trong 4 thuật toán.

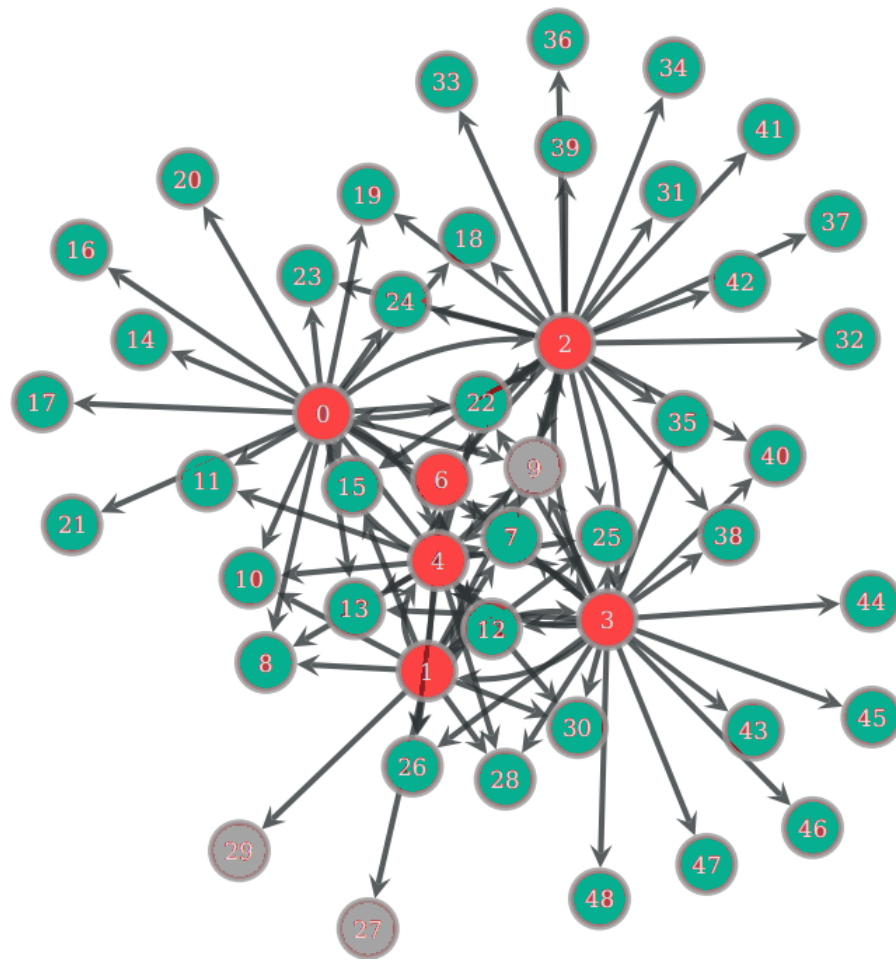
Về mặt thời gian thực hiện, STMB cũng thể hiện tốc độ vượt trội của mình so

với các thuật toán được đánh giá là thời gian thực thi thấp như Page Rank hay High-Degree. Do thời gian thực hiện của thuật toán Greedy với dữ liệu thu thập được trung bình hơn 400(s) lâu hơn rất nhiều so với các thuật toán khác nên chúng tôi không đưa ra trong hình 4.6.



Hình 4.6: So sánh thời gian thực hiện của các thuật toán với dữ liệu thu thập được cho bài toán TMB

Tiếp tục sử dụng công cụ graph\_tool chúng tôi mô phỏng lại phân trung tâm của đồ thị và thu được kết quả như trong hình 4.7.



Hình 4.7: Kết quả mô phỏng với dữ liệu thu thập được trong trường hợp  $\gamma = 15$  của thuật toán STMB

Với các đỉnh được tô màu đỏ thể hiện nguồn phát tán thông tin sai lệch, các đỉnh được tô màu xám là tập đỉnh cần loại bỏ ra khỏi đồ thị, các đỉnh còn lại được tô màu xanh.

Từ những kết quả trên, kết hợp với những kết quả đã được chứng minh về mặt lý thuyết, có thể thấy được hiệu quả của hai giải pháp nhóm tác giả đề xuất. Việc kiểm chứng dựa trên cả các bộ dữ liệu thực, và các dữ liệu thu thập trực tiếp từ mạng xã hội Facebook đảm bảo các giải pháp đưa ra vừa thỏa mãn được tính chất học thuật cao, vừa thỏa mãn được khả năng ứng dụng thực tế ngay tại Việt Nam.

#### 4. Ảnh xạ người dùng thực

Với kết quả thu được của mỗi giải pháp tương ứng với từng kịch bản là danh sách các đỉnh có vai trò quang trọng đối với quá trình phát tán thông tin sai lệch, nhóm tác giả tiến hành dựa vào tệp tin ID.txt để tìm được tài khoản Facebook những người dùng đóng vai trò quan trọng đây.

- Đối với giải pháp FLE, kết quả thu được thể hiện trong bảng 4.7

ID	Tên tài khoản	Số lượng người theo dõi
100002788524244	Chú Tễu	43.974
634032152	Đinh Công Bằng	non
100000315904157	Trịnh Anh Tuấn	13.611
1849592060	Thoai Huu Dinh	20.719
100004460193964	Nguyễn Duy Cường	64.779
100003888766757	Trương Thị Hà	11.320
100000757109103	Nguyen Thien Nhan	19.460
100001730194864	Ngọc Hà	629
100005228346422	Nguyễn Hồng Quân	non
100003276218628	Trần Lai	non

Bảng 4.7: Danh sách người dùng thực - Kết quả giải pháp FLE

- Đối với giải pháp STMB, kết quả thu được thể hiện trong bảng 4.8.

ID	Tên tài khoản	Số lượng người theo dõi
1674099665	Lê Nguyễn Hương Trà	468.761
1763201593	Huynh Ngoc Chenh	78.989

Bảng 4.8: Danh sách người dùng thực - Kết quả giải pháp SMTB

Kết quả thu được có thể thấy là những người dùng có số lượng theo dõi trên mạng xã hội Facebook lớn, là đối tượng tiềm năng, có vai trò quan trọng đối với việc phát tán thông tin sai lệch. Đây là tập những người dùng nhóm tác giả đề xuất sử dụng các giải pháp ngăn chặn, các biện pháp vô hiệu hóa, đặt giám sát, thuyết phục đối tượng để đối tượng không còn khả năng lan truyền thông tin sai lệch.

## 5. Giải pháp ngăn chặn

Tuy các giải pháp của nhóm tác giả đã đề xuất ra được danh sách những người dùng đóng vai trò quan trọng trong quá trình lan truyền thông tin sai lệch, tương ứng với hai kịch bản có ứng dụng cao trong thực tế. Tuy nhiên trong quá trình nghiên cứu, nhóm tác giả nhận thấy việc áp dụng biện pháp ngăn chặn với các tài khoản này còn gặp phải một số vấn đề khó khăn.

Theo Điều 22, Luật An ninh Quốc gia nước Cộng hòa xã hội chủ nghĩa Việt Nam năm 2004, Công An nhân dân mà một trong các cơ quan chuyên trách được giao thực

hiện nhiệm vụ bảo vệ An ninh quốc gia. Theo Điều 10, Luật Viễn thông năm 2009 quy định Cơ quan quản lý chuyên ngành về viễn thông là cơ quan thuộc Bộ Thông tin và Truyền thông. Tuy nhiên trong thực tế, các cơ quan chức năng của Nhà nước ta chưa có quyền quản lý đối với các nội dung trên mạng xã hội, được quản lý bởi các tập đoàn Facebook, Twitter, Microsoft. Do đó, việc xóa bỏ các nội dung, thông tin sai lệch lan truyền trên mạng xã hội, hay vô hiệu hóa người dùng đăng tải thông tin sai lệch là chưa khả thi.

Hiện nay một số nước trên thế giới đã và đang thành lập các trung tâm chống tin tức giả, trong đó có thể kể đến một số giải pháp đã được áp dụng như sau:

- Xóa bỏ các tài khoản có vai trò quan trọng trong lan truyền thông tin xấu.
  - Năm 2016, Twitter đã loại bỏ 125.000 tài khoản được nghi ngờ có liên quan đến các hoạt động khủng bố.
  - Năm 2017, Facebook đã xóa 30.000 tài khoản giả mạo lan truyền các tin đồn trước cuộc bầu cử tổng thống Pháp năm 2017.
- Thuyết phục tài khoản không tin, lan truyền thông tin xấu. Biện pháp này được đề cập trong nghiên cứu của Zhang và các cộng sự [50, 51].
- Khử nhiễm bằng cách sử dụng các tin tốt tác động họ. Biện pháp này được Nam P.Nguyen và các cộng sự sử dụng trong nghiên cứu [2].
- Ngoài ra có thể sử dụng các biện pháp kinh tế, tài chính để phổ biến, nâng cao ý thức cho những người dùng trong danh sách này.

Đặc biệt trong hoàn cảnh dự thảo Luật An ninh mạng đã được thông qua, Luật An ninh mạng đang được chỉnh lý và sắp chính thức có hiệu lực. Vai trò quản lý của Nhà nước, các cơ quan chuyên trách đối với các nguồn thông tin trên mạng ngày càng được nâng cao. Việc phát hiện và ngăn chặn thông tin sai lệch đang trở thành một vấn đề cấp bách, có ý nghĩa thiết thực.

## KẾT LUẬN

Sự bùng nổ của thông tin sai lệch trên các mạng xã hội đang là một nguy cơ lớn đối với người dùng. Sự lan truyền nhanh chóng của thông tin sai lệch đến từ sự chủ quan của người dùng hoặc được lan truyền một cách có chủ đích với động cơ, mục đích không trong sáng. Hơn nữa, hoạt động này đem lại những hậu quả nghiêm trọng đối với cá nhân, tổ chức không chỉ về kinh tế chính trị mà còn tác động đến tâm lý, cuộc sống con người, gây mất ổn định an ninh quốc gia. Do đó, việc đưa ra giải pháp hiệu quả nhằm ngăn chặn kịp thời sự lan truyền của thông tin sai lệch trên mạng xã hội trực tuyến là việc làm hết sức cấp thiết nhằm giảm thiểu tối đa thiệt hại do thông tin sai lệch gây ra đối với cá nhân, tổ chức, góp phần làm trong sạch môi trường mạng, đồng thời nâng cao sự tin tưởng của người dùng đối với những thông tin trên mạng xã hội.

Trong đề tài này, nhóm tác giả đề xuất hai kịch bản ngăn chặn có tính ứng dụng trong thực tế, tương ứng là hai bài toán tối ưu. Dựa trên 2 mô hình lan truyền thông tin thường được sử dụng là mô hình ngưỡng tuyến tính (Linear Threshold Model), mô hình bậc độc lập (Independent Cascade Model), nhóm tác giả đề xuất mô hình ngưỡng tuyến tính với bước thời gian rời rạc (Time – Discrete Linear Threshold Model) từ đó đưa ra hai giải pháp ngăn chặn hiệu quả. Đề tài đã đạt được một số kết quả chính như sau:

- Tìm hiểu tổng quan về mạng xã hội, các đặc trưng cơ bản của mạng xã hội, lợi ích của mạng xã hội, các nguy cơ lây lan và tác hại của thông tin sai lệch trên mạng xã hội.
- Tìm hiểu cơ chế lan truyền thông tin và đặc tính của hai mô hình lan truyền thông tin thường được sử dụng, từ đó đề xuất một mô hình lan truyền mới, phù hợp với bài toán khảo sát.
- Đề xuất hai kịch bản, hai bài toán tương ứng là LSE và TMB, chứng minh hai bài toán này thuộc lớp #P – Khó, đề xuất hai thuật toán hiệu quả FLE và STMB với tốc độ thực thi nhanh và hiệu quả tốt.
- Thu thập dữ liệu thực từ mạng xã hội Facebook với phạm vi xung quanh các đối tượng cầm đầu, chủ mưu các tổ chức phản động, đối tượng là người có uy tín lớn. Từ đó áp dụng giải pháp FLE và STMB để đưa ra tập người dùng tiềm năng để vô hiệu hóa hoạt động của họ và đem lại hiệu quả tốt nhất.

### Hướng phát triển - kiến nghị

Trong thời gian tới, nhóm tác giả đề xuất một số hướng phát triển của đề tài như sau:



- Thiết kế thuật toán đạt tỉ lệ tối ưu cao hơn, thời gian chạy thuật toán nhanh hơn cho bài toán hạn chế ảnh hưởng của thông tin sai lệch trên mạng xã hội.
- Mở rộng bài toán hạn chế thông tin sai lệch đối với nhiều nguồn thông tin sai lệch cùng tác động lên mạng.
- Thiết lập một phần mềm với giao diện trực quan đóng gói tất cả các công cụ để cho người dùng có thể sử dụng dễ dàng.
- Tìm hiểu các giải pháp thu thập dữ liệu hiệu quả hơn, thực thi với tốc độ nhanh hơn.
- Sau khi Luật An ninh mạng được thông qua, đề xuất các cơ quan chức năng nghiên cứu, mở rộng và áp dụng những thành tựu, kết quả, giải pháp mà nhóm tác giả đã đề xuất.

## TÀI LIỆU THAM KHẢO

- [1] Bộ Thông tin và Truyền thông đưa ra biện pháp ngăn chặn thông tin xấu độc trên MXH. Truy xuất từ: <https://news.zing.vn/thong-tin-tieu-cuc-chu-yeu-tren-mang-xa-hoi-cua-nuoc-ngoai-post796383>.
- [2] Nam P. Nguyen, Guanhua Yan, My T. Thai (2013), "Analysis of misinformation containment in online social networks", Computer Networks.
- [3] ] N.A Christakis, J. H. Fowler (2010), "Social network sensors for early detection of contagious outbreak", PloS one. doi:10.1371/journal.pone. 0012948.
- [4] Elias Boutros Khalil, Bistra Dilkina, and Le Song. 2014. Scalable diffusion-aware optimization of network topology. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. New York, NY, USA, 1226–1235.
- [5] Ceren Budak, Divyakant Agrawal, Amr El Abbadi, Limiting the Spread of Misinformation in Social Networks.
- [6] CV Pham, MT Thai, HV Duong, BQ Bui, HX Hoang. Maximizing misinformation restriction within time and budget constraints. Journal of Combinatorial Optimization, 1-39.
- [7] P.Cui, S. Jin, L.Yu, F.Wang, W.Zhu, S. Yang (2013), " Cascading outbreak prediction in network: a data-driven approach", In: Proceeding of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, Chicago, pp. 901-909.
- [8] Quy định việc cung cấp, sử dụng Internet của chính phủ. Truy xuất từ: <https://thuvienphapluat.vn/van-ban/Cong-nghe-thong-tin/Nghi-dinh-72-2013-ND-CP-quan-ly-cung-cap-su-dung-dich-vu-Internet-va-thong-tin-tren-mang-201110.aspx>.
- [9] Karlova, Natascha, Fisher Karen. A Social Diffusion Model of Misinformation and Disinformation for Understanding Human Information Behaviour. Information Research, 2013, pp. 1-17.
- [10] Tin tặc giả mạo đăng tin sai lệch trên hăng thông tấn Associated Press. Truy xuất từ: <https://www.usatoday.com/story/theoval/2013/04/23/obama-carney-associated-press-hack-white-house/2106757>

- [11] Ảnh hưởng của MXH với bầu cử Pháp. Truy xuất từ: <https://baomoi.com/facebook-va-phep-thu-mang-ten-bau-cu-phap/c/22045463.epi>.
- [12] Tin đồn dịch Ebola bùng phát tại Hà Nội. Truy xuất từ: <http://baophapluat.vn/tin-nong/lam-ro-danh-tinh-2-nguoi-tung-tin-don-that-thiet-ve-benh-dich-ebola-194152.html>
- [13] Tin đồn công ty Formosa xả thải chưa qua xử lí. Truy xuất từ: <http://m.vietnamnet.vn/vn/thoi-su/moi-truong/bo-tn-mt-bac-bo-tin-don-formosa-phat-thai-dioxin-370436.html>
- [14] H Zhang, M. Alim, M. T. Thai, and H. Nguyen, Monitor Placement to Timely Detect Misinformation in Online Social Networks, in Proceedings of the 2015 IEEE International Conference on Communications (ICC), 2015.
- [15] D. Kempe, J. Kleinberg, and E. Tardos. 2003. Maximizing the spread of influence through a social network. In Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. Washington, DC, USA, 137-146. <https://doi.org/10.1145/956750.956769>.
- [16] D. Kempe, J. Kleinberg, and E. Tardos. Influential Nodes in a Diffusion Model for Social Networks. In ICALP, 2005, pp. 1127-1138.
- [17] J. Goldenberg, B. Libai, and E. Muller. Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth. Marketing Letters, 2001, pp. 211-223
- [18] T. Carnes, R. Nagarajan, S. M. Wild, and A. V. Zuylen. Maximizing Influence in a Competitive Social Network: a Follower's Perspective. In Proceedings of the Ninth International Conference on Electronic Commerce, 2007, pp. 351-360.
- [19] J. Leskovec, M. Mcglohon, C. Faloutsos, N. Glance, and M. Hurst. Cascading Behavior in Large Blog Graphs. In Proceedings of the 2007 SIAM International Conference on Data Mining, 2007, pp. 551-556.
- [20] Pedro Domingos and Matthew Richardson. Mining the Network Value of Customers. In Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2001, pp. 57-66.
- [21] A. Goyal, W. Lu, L. V.S. Lakshmanan. Celf++: Optimizing the Greedy Algorithm for Onfluence Maximization in Social Networks. In Proc. WWW, 2011, pp. 47-48.

- [22] A. Goyal, F. Bonchi, L. V. S. Lakshmanan, and S. Venkatasubramanian.: On minimizing budget and time in influence propagation over social networks. *Social Network Analysis and Mining*, 2(1), 179–192 (2012)
- [23] Wei Chen, Laks V.S. Lakshmanan, and Carlos Castillo. *Information and Influence Propagation in Social Networks*. A Publication in the Morgan & Claypool Publishers series Synthesis Lectures on Data Management, 2014.
- [24] V. Qazvinian, E. Rosengren, D. R. Radev, and Qiaozhu Mei. Rumor has it: Identifying Misinformation in Microblogs. In *Proc. EMNLP*, 2011, pp. 1589-1599.
- [25] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang. 2013. Prominent Features of Rumor Propagation in Online Social Media. In *Proc. ICDM*, 2013, pp. 1103-1108.
- [26] D. T. Nguyen, N. P. Nguyen, and M. T. Thai. Sources of Misinformation in Online Social Networks: Who to Suspect. In *Proceedings of the IEEE Military Communications Conference*, 2012.
- [27] D. T. Nguyen, N. P. Nguyen, and M. T. Thai. Sources of Misinformation in Online Social Networks: Who to Suspect. In *Proceedings of the IEEE Military Communications Conference*, 2012.
- [28] Ceren Budak, Divyakant Agrawal, Amr El Abbadi, Limiting the Spread of Misinformation in Social Networks.
- [29] Huiling Zhang, Md Abdul Alim, Xiang Li, My T. Thai, and Hien T. Nguyen. 2016. Misinformation in online social networks: Detect them all with a limited budget. *ACM Trans. Inf. Syst.* 34, 3, Article 18 (April 2016), 24 pages. DOI:<http://dx.doi.org/10.1145/2885494>.
- [30] Zhang Y, Prakash B (2015) Data-aware vaccine allocation over large networks. *ACM Trans Knowl Discov Data*. <https://doi.org/10.1145/2803176>.
- [31] Zhang H, Alim M, Li X, My TT, Nguyen H (2016a) Misinformation in online social networks: catch them all with limited budget. *ACM Trans Inf Syst* 34(3):18.
- [32] Zhang Y, Adigay A, Saha S, Vullikanti A, Prakash A (2016b) Near-optimal algorithms for controlling propagation at group scale on networks. *IEEE Trans Knowl Data Eng*. <https://doi.org/10.1109/TKDE.2016.2605088>

- [33] Yadron D (2017) Twitter deletes 125,000 Isis accounts and expands anti-terror teams. The Guardian Web. <https://www.theguardian.com/technology/2016/feb/05/twitter-deletes-isis-accounts-terrorismonline>. Accessed 24 June 2017.
- [34] Kottasov I (2017) Facebook targets 30,000 fake accounts in France. CNN media Web. <http://money.cnn.com/2017/04/14/media/facebook-fake-news-france-election/index.html>. Accessed 24 June 2017.
- [35] Zaixin Lu, Wei Zhang, Weili Wu, Joonmo Kim, and Bin Fu. 2011. The complexity of influence maximization problem in the deterministic linear threshold model. *Journal of Combinatorial Optimization* 24, 3 (April 2011), 374–378. <https://doi.org/10.1007/s10878-011-9393-3>.
- [36] Vijay V. Vazirani. 2001. *Approximation Algorithms*. Springer, Verlag New York, Inc. New York, NY, USA.
- [37] L. G. Valiant.: The complexity of enumeration and reliability problems. *SIAM Journal on Computing* 8(3), 410–421 (1979)
- [38] DT Nguyen, H Zhang, S Das, MT Thai, TN Dinh.: Least cost influence in multiplex social networks: Model representation and analysis. In: *Proceedings of IEEE 13th International Conference on Data mining*, pp. 567–576, IEEE(2013)
- [39] J. Leskovec, J. Kleinberg and C. Faloutsos. *Graphs over Time.: Densification Laws, Shrinking Diameters and Possible Explanations*. In: *Proceedings of the 11th ACM SIGKDD international conference on Knowledge discovery in data mining*, pp. 177–187, ACM (2005)
- [40] W. Chen, C. Wang, and Y. Wang.: *Scalable Influence Maximization in Social Networks under the Linear Threshold Model*. In: *Proceedings of the 2010 IEEE International Conference on Data Mining*, pp. 88–97, IEEE(2010)
- [41] Elias Boutros Khalil, Bistra Dilkina, Le Song.: *Scalable Diffusion–Aware Optimization of Network Topology*. In: *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1226–1235, ACM (2014)
- [42] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. M. Van–Briesen, and N. S. Glance.: *Cost–effective outbreak detection in networks*. In: *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, 420–429, ACM (2007).

- [43] H. Tong, B. A. Prakash, C. Tsourakakis, T. Eliassi-Rad, C. Faloutsos, and D. H. Chau.: On the vulnerability of large graphs. In: Proceedings of IEEE International Conference on Data Mining (ICDM), pp. 1091–1096, IEEE (2010)
- [44] Prakash BA, Tong H, Valler N, Faloutsos M, Faloutsos C (2010).: Virus propagation on time-varying networks: Theory and immunization algorithms. In: Proceedings of 2010 Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 99–114, Springer (2010)
- [45] Y. Zhang, A. Adiga, S. Saha, A. Vullikanti, and B. A. Prakash.: Nearoptimal algorithms for controlling propagation at group scale on networks. *IEEE Trans. Knowl. Data Eng.*, 28(12), 3339–3352 (2016)
- [46] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. M. Van-Briesen, and N. S. Glance.: Cost-effective outbreak detection in networks. In: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 420-429, ACM(2007)
- [47] Amit Goyal, Wei Lu, and Laks V.S. Lakshmanan. 2011. SIMPATH: An Efficient Algorithm for Influence Maximization under the Linear Threshold Model. In Proceeding IEEE 11th International Conference on Data Mining. Vancouver, Canada, 211-220. <https://doi.org/10.1109/ICDM2011.132>.
- [48] M. Cha, A. Mislove, and K.P. Gummadd. 2009. A measurement-driven analysis of information propagation in the Flickr social network. In Proceeding of the 18th international conference on World Wide Web (2009). ACM, Madrid, Spain, 721–730. <https://doi.org/10.1145/1526709.1526806>.
- [49] J. Leskovec, L. A. Adamic, and B. A. Huberman. 2007. The dynamics of viral marketing. *ACM Transaction on the Web* 1, 1 (May 2007). <https://doi.org/10.1145/1232722.1232727>.
- [50] Zhang Y, Prakash B (2015) Data-aware vaccine allocation over large networks. *ACM Trans Knowl Discov Data*. <https://doi.org/10.1145/2803176>
- [51] Zhang H, Alim M, Li X, My TT, Nguyen H (2016a) Misinformation in online social networks: catch them all with limited budget. *ACM Trans Inf System*.