# **3. Các nghiên cứu liên quan**

Chúng tôi điểm lại các nghiên cứu hiện tại trên các phương pháp model-based CF, graph-based CF và graph neural network-based, là các nghiên cứu liên quan chặt chẽ đến bài báo này. Từ đó, chúng tôi làm rõ điểm khác biệt của các phương pháp này với phương pháp NGCF chúng tôi đề xuất.

## **3.1. Phương pháp Model-based CF**

Các recommendation system hiện đại [5, 14, 33] tham số hóa user và item bằng cách biểu diễn chúng dưới dạng vector và xây dựng lại dữ liệu tương tác giữa user-item dựa trên các tham số mô hình. Ví dụ, MF [20, 26] chiếu ID của mỗi user và item thành 1 embedding vector, và tính tích vô hướng của chúng để dự đoán tương tác. Để làm tốt nhiệm vụ embedding, người ta đã cố gắng tích hợp nhiều thông tin như nội dung item [2, 30], quan hệ xã hội [34], quan hệ item [37], user review [3], và các external knowledge graph [32, 35]. Mặc dù tích vô hướng có thể thể hiện quan hệ user-item của 1 tương tác đã biết, tính tuyến tính của nó là rào cản trong việc thể hiện các mối quan hệ phức tạp, không tuyến tính giữa user-item [14, 15]. Vì vấn đề này, các nghiên cứu gần đây [11, 14, 15, 36] tập trung vào công nghệ học sâu để thể hiện tốt hơn tương tác không tuyến tính giữa user và item. Ví dụ, mô hình neural CF, như NeuMF [14], triển khai mạng neural để thể hiện quan hệ tương tác, trong khi đó, mô hình translation-based CF, như LRML [28] mô hình hóa quan hệ tương tác bằng độ đo khoảng cách Euclid.

Mặc dù đã đạt được nhiều thành công, nhưng chúng tôi vẫn cho rằng thiết kế phương thức embedding là không hiệu quả để có được kết quả embedding tối ưu cho CF vì đó là thông tin ẩn. Nhìn chung, các phương pháp nói trên công việc embedding là biến đổi các đặc trưng mô tả (ví dụ ID, các thuộc tính) thành dạng vector, và tương tác được thể hiện là độ tương tự giữa các vector. Khi tương tác user-item được xây dựng lại toàn bộ, các thuộc tính dẫn xuất của tính tương đồng hành vi cũng sẽ được ghi lại. Tuy nhiên, các hiệu ứng dẫn xuất này không được mã hóa tường minh, vì vậy không có gì đảm bảo rằng các kết nối không trực tiếp giữa user và item là đóng trong không gian embedding. Không có mã hóa tường minh của thông tin CF, sẽ khó để có được kết quả embedding đáp ứng các thuộc tính yêu cầu.

## **3.2. Phương pháp Graph-based CF**

Một hướng nghiên cứu khác [12, 24, 40] khai thác đồ thị tương tác user-item để suy luận ra các item ưa thích của user. Các nghiên cứu ban đầu như ItemRank [7] và BiRank [12] dùng ý tưởng lan truyền nhãn để ghi lại hiệu ứng CF. Để chấm điểm các item cho 1 user, các phương pháp này xác định các nhãn của các item tương tác của user, sau đó lan truyền các nhãn này trên đồ thị. Vì điểm khuyến nghị được tính toán dựa trên tham số tầm với (reachness) (gần giống với tham số độ tương tự) giữa các item trong quá khứ và các item khuyến nghị, các phương pháp này gọi là nhóm neighbor-based. Tuy nhiên, các phương pháp này về ý tưởng cũng thuộc nhóm model-based CF vì nó thiếu các tham số mô hình để tối ưu hàm mục tiêu của hệ khuyến nghị.

Nghiên cứu gần đây là HOP-Rec [40] khắc phục vấn đề trên bằng cách kết hợp 2 phương pháp graph-based và embedding-based. Đầu tiên, nó thực hiện random walk để làm giàu tương tác của 1 user với các item kết nối đến nó. Sau đó, nó huấn luyện MF bằng BPR dựa trên dữ liệu tương tác user-item đã được làm giàu để xây dựng mô hình khuyến nghị. Hiệu quả của phương pháp HOP-Rec thông qua phân rã ma trận cung cấp bằng chứng cho thấy kết hợp các thông tin kết nối là sẽ cho các bembedding tốt hơn để ghi lại các hiệu ứng CF. Tuy nhiên, chúng tôi cho rằng HOP-Rec không tìm được đầy đủ các kết nối bậc cao mà chỉ được sử dụng để làm giàu dữ liệu huấn luyện, thay vì được sử dụng trực tiếp cho việc embedding của mô hình. Hơn nữa, hiệu năng của HOP-Rec phụ thuộc nhiều vào random walk, vì vậy đòi hỏi được tinh chỉnh cẩn thận, ví dụ như một thiết lập phù hợp cho hệ số phân rã (decay factor).

## **3.3. Mạng tích chập đồ thị (Graph Convolutional Network)**

Bằng cách nghĩ ra phương pháp xử lý tích cập đồ thị chuyên biệt trên đồ thị tương tác user-item, (cf. Equation 3), chúng tôi làm cho NGCF hiệu quả trong phát hiện các thông tin CF trên các kết nối bậc cao. Ở đây, chúng tôi bàn luận về các phương pháp khuyến nghị hiện có đang triển khai xử lý tích chập đồ thị [29, 42, 43].

GC-MC [29] ứng dụng mạng tích chập đồ thị (GCN) [18] trên đồ thị user-item, tuy nhiên nó chỉ triển khai 1 lớp tích chập để phát hiện ra các kết nối trực tiếp giữa user và item. Vì vậy, nó không lấy được thông tin kết nối mức cao của tương tác. PinSage [42] là 1 giải pháp công nghiệp trong đó triển khai nhiều lớp tích chập đồ thị lên đồ thị item-item cho hệ thống đề xuất ảnh Pinterest. Với hệ thoogns này, hiệu ứng CF được ghi lại trên mức quan hệ item, hơn là mức thu thập hành vi người dùng. SpectralCF [43] đề xuất xử lý tích chập phổ để phát hiện tất cả các kết nối có thể có giữa các user và các item trong vùng phổ. Thông qua phân tích trị riêng của ma trận kề của đồ thị, có thể phát hiện được các kết nối giữa các cặp user-item. Tuy nhiên, phân tích trị riêng có độ phức tạp tính toán cao, đòi hỏi nhiều thời gian tính toán và khó khăn khi triển khai các hệ thống đề xuất lớn.

# **4. Thực nghiệm**

Chúng tôi tiến hành thực nghiệm trên 3 tập dữ liệu thực tế để đánh giá phương pháp đã đề xuất, đặc biệt là lớp lan truyền embedding. Mục tiêu của chúng tôi là trả lời các câu hỏi sau:

**RQ1:** So sánh hiệu năng của NGCF với các phương pháp CF tiên tiến nhất hiện nay?

**RQ2:** Tác động của các bộ tham số khác nhau (ví dụ về chiều sâu của mạng, lớp lan truyền embedding, cơ chế tổng hợp các lớp, message dropout, node dropout)?

**RQ3:** Kết nối bậc cao tác động lên kết quả trình diễn như thế nào?

## **4.1. Mô tả tập dữ liệu**

Để đánh giá tính hiệu của của NGCF, chúng tôi tiến hành thực nghiệm trên 3 tập dữ liệu công khai: Gowalla, Yelp2018 và Amazon-book, có sự khác biệt về lĩnh vực, kích thước, và độ thưa. Đặc điểm thống kê của 3 tập dữ liệu này trong bảng 1

A picture containing text, screenshot, font, number

Description automatically generated

**Gowalla**: Đây là dữ liệu chứa thông tin địa điểm của người dùng được họ chia sẻ khi check in. Để đảm bảo chất lượng dữ liệu, người ta chỉ lấy dữ liệu user, item mà có ít nhất 10 tương tác.

**Yelp2018**: Dữ liệu thu thập từ thử thách Yelp 2018, trong đó các item là các nhà hàng, quán bar địa phương. Người ta cũng chỉ lấy các item có ít nhất 10 order để đảm bảo chất lượng dữ liệu.

**Amazon-book**: các review của Amazon là tập dữ liệu được sử dụng rộng rãi cho hệ gợi ý sản phẩm [9]. Nhóm tác giả chọn Amazon-book từ tập dữ liệu đó. Họ cũng chỉ sử dụng các user, item có ít nhất 10 tương tác.

Với mỗi tập dữ liệu, nhóm tác giải chọn ngẫu nhiên 80% tương tác quá khứ cho mỗi user để xây dựng tập dữ liệu training, phần còn lại làm tập dữ liệu test. Trong tập dữ liệu training, tiếp tục chọn ngẫu nhiên 10$ các tương tác làm tập validation để tinh chỉnh các siêu tham số. Với mỗi tương tác user-item quan sát được, tác giả coi đó là trường hợp dương tính, và sau đó xây dựng chiến lược lấy mẫu âm tính để đối sánh mẫu dương tính với mẫu âm tính mà khách hàng không mua trước đó.

## **4.2. Thiết lập thí nghiệm**

### **4.2.1. Độ đo đánh giá**

Với mỗi user trong tập dữ liệu test, người ta coi tất cả các item mà user không tương tác trước đó là item âm tính. Sau đó, các phương pháp cho ra điểm tham chiếu của mỗi user đối với tất cả các item, ngoại trừ các item dương tính trong tập training. Để đánh giá tính hiệu quả của top K gợi ý và xếp hạng tham chiếu, người ta dùng 2 giao thức đánh giá phổ biến [14, 40] là recall@K và ndcg@K3. Mặc định, tác giả đặt K=20. Họ báo cáo kết quả đo lường trung bình với tất cả các user trong tập test.

### **4.2.2. Baseline**

Để chứng minh tính hiệu quả, tác giả so sánh phương pháp NGCF với các phương pháp sau:

**MF** [26]: đây là phương pháp phân rã ma trận được tối ưu bằng hmaf lost BPR (Bayesian personalized ranking), trong dó phát hiện được cá tương tác trực tiếp user-item chỉ như giá rị mục tiêu của hàm tương tác.

**NeuMF** [14]: đây là phương pháp tiên tiến của mô hình neural CF, trong đó sử dụng nhiều lớp ẩn phía trước các tính toán element-wise và ghép nối (concaternation) giữa embedding của user và item để bắt được các tương tác phi tuyến giữa chúng. Đặc biệt, tác giả triển khai kiến trúc 2 lớp, trong đó kích thước của mỗi lớp ẩn giống nhau.

**CMN** [5]: đây là phương pháp tiên tiến của mô hình memory-based, trong đó đại diện user kết hợp các khe nhớ của các user hàng xóm thông qua các memory layer. Lưu ý rằng các kết nối bậc nhất được sử dụng để tìm các user tương tự đã tương tác với cùng item.

**HOP-Rec** [40]: đây là phương pháp tiên tiến của mô hình graph-based, trong đó các hàng xóm bậc cao thu được từ random walks được phát hiện để làm giàu dữ liệu tương tác user-item.

**PinSage** [42]: PinSage được thiết kế để triển khai GraphSAGE [8] trên đồ thị item-item. Trong thực nghiệm này, tác giả chỉ áp dụng nó trên đồ thị tương tác user-item. Đặc biệt, tác giả triển khai 2 lớp tích chập đồ thị như khuyến nghị tại [42], và kích thước các lớp ẩn này được đặt bằng với kích thước embedding.

**GC-MC** [29]: Mô hình này sử dụng phần mã hóa GCN [18] để tạo ra các đại diện cho user, item, trong đó các hàng xóm bậc 1 được xem xét. Vì vậy một lớp tích chập đồ thị, trong đó kích thước lớp ẩn được đặt bằng kích thước embedding như khuyến nghị trong [29].

Tác giả cũng thử SpectralCF [43] nhưng thấy rằng phân tích trị riêng tốn nhiều chi phí thời gian và tài nguyên, đặc biệt khi số lượng user và items lớn. Vì vậy, mặc dù phương pháp này có hiêu quả tốt với tập dữ liệu nhỏ, tác giả đã không chọn nó để so sánh. Để đảm bảo so sánh khách quan, tất cả cac sphwong pháp đều sử dụng hàm losst BPR như trong công thức (11).

### **4.2.3. Thiết lập tham số**

Tác giả triển khai NGCF trên Tensorflow. Kích thước embedding đặt cố định bằng 64 cho tất cả các mô hình. Với HOP-Rec, tác giả tìm bước random walk trong {1, 2, 3} và tinh chỉnh learning rate trong {0.025, 0.020, 0.015, 0.010}. Tác giả tối ưu tất cả các mô hình ngoại trừ HOP-rec bằng Adam optimizer. Trong đó batch size cố định là 1024. Tác giả dùng tìm kiếm lưới để tìm các siêu tham số: learning rate thuộc tập {0.0001, 0.0005, 0.001, 0.005}, hệ số của chuẩn hóa L2 thuộc được tìm kiếm trong , tỷ lệ dropout thuộc . Bên cạnh đó, tác giả triern khai kỹ thuật dropout node cho GC-MC và NGCF, trong đó tỷ lệ dropout thuộc . Tác giả sử dụng khởi tạo Xavier để khởi tạo các tham số của mô hình. Ngoài ra, tác giả cũng triển khai chiến thuật early stopping, ví dụ sẽ dừng nếu điểm recall@20 trên tập validataion không tăng sau 50 epoch. Để mô hình thông tin mã hóa của CF trên kết nối bậc 3, tác giả đặt độ sâu của NGCF L là 3. Không cần đặc tả, tác giả chỉ ra kết qủa của 3 lớp lan truyền embedding, tỷ lệ dropout node là 0, dropout message là 0.1

## **4.3. So sánh hiệu năng (RQ1)**

Tác giả bắt đầu bằng việc so sánh hiệu năng của tất cả các phương pháp, sau đó kiểm tra tính ưu việt của việc mô hình hóa các kết nối bậc cao trong điều kiện cấu hình thưa.

### **4.3.1. So sánh tổng thể**

Trong bảng 2 là kết quả so sánh hiệu năng.

A picture containing text, screenshot, font, number

Description automatically generated

Có thể thấy:

* MF có hiệu năng kém trên cả 3 tập dữ liệu. Điều này cho thấy tích vô hướng không hiệu quả trong việc bắt được các quan hệ phức tạp giữa user và item, sau đó hạn chế về hiệu năng. NeuMF có hiệu năng tốt hơn MF trên tất cả các trường hợp, chứng tỏ tầm quan trọng của các tương tác phi tuyến giữa user và item. Tuy nhiên, cả MF và NeuMF đều mô hình hóa rõ ràng kết nối trong quá trình học embedding, trong đó dễ dàng dẫn tới biểu diễn cận tối ưu.
* So với MF và NeuMF, hiệu năng của GC-MC chứng tỏ rằng việc hợp nhất các hàng xóm bậc 1 có thể cải thiện việc học. Tuy nhiên, trên tập Yelp2018\*, GC-MC có hiệu năng thấp hơn NeuMF với điểm ndcg@20. Lý do có thể là GC-MC không phát hiện được các tương tác phi tuyến giữa user và item.
* CMN nhìn chung có hiệu năng tốt hơn GC-MC trong hầu hết trường hợp. Sự cải thiện này có thể là do cơ chế neural attention, trong đó xác định được trọng số chú ý của mỗi user hàng xóm, thay vì dùng các weigh bằng nhau hoặc weight heuristic trong GC-MC.
* PinSage có hiệu năng thấp hơn 1 chút so với CMN trên tập Gowalla và Amazon-Book, nhưng lại có hiệu năng tốt hơn rất nhiều trên tập Yelp2018\*; trong khi đó, HOP-Rec nói chung có hiệu năng tốt hơn đáng kể trong hầu hết các trường hợp. Điều này rất có ý nghĩa vì PinSage giới theieuj kết nối bậc cao trong hàm embedding, và HOP-Rec phát hiện được các hàng xóm bậc cao để làm giàu dữ liệu training, trong khi CMN chỉ quan tâm đến các user giống nhau. Việc này cho thấy mặt tích cực của việc mô hình hóa các kết nối và hàng xóm bậc cao.
* NGCF thể hiện hiệu năng tốt nhất trên tất cả các tập dữ liệu. Cụ thể, NGCF đã vượt qua được mức baseline tốt nhất về điểm recall@20 lần lượt 11.68%, 11.97%, 9.61% trên các tập dữ liệu Gowalla, Yelp2018\* và Amazon-Book. Bằng cách xếp chồng các lớp lan truyền embedding, NGCF có thể phát hiện được các kết nối bậc cao 1 cách rõ ràng, trong khi CMN và GC-MC chỉ có thể phát hiện được các hàng xóm bậc 1 để dẫn dắt quá trình học. Điều này chứng minh tầm quan trọng của việc bắt được các thông tin tương tác trong nhiệm vụ embedding. Hơn nữa, so với PinSage, NGCF xem xét nhiều khía cạnh để suy luận ra nội dung ưa thích của user, trong khi PinSage chỉ sử dụng đầu ra của lớp cuối. Điều này cho thấy các lớp lan truyền khác nhau mã hóa các thông tin khác nhau. Và hiệu năng đạt được cao hơn so với HOP-Rec cho thấy mã hóa CF mã hóa tường minh trong khâu embedding mang lại hiệu quả tốt hơn. Tác giả triển khai kiểm định t-test 1 mẫu và p-value<0.05 chứng tỏ hiệu quả vượt trội của NGCF so với baseline tốt nhất (các trường hợp gạch chân trong bảng 2) là có ý nghĩa thống kê.

### **4.3.2. Performance Comparison w.r.t Interaction Sparsity Levels**

Tính thưa thường là nguyên nhân gây hạn chế hiệu quả của hệ thống khuyến nghị, vì 1 số ít tương tác của các user inactive không đủ để tạo ra các biểu diễn chất lượng cao. Tác giả nghiên cứu xem liệu việc phát hiện ra các thông tin kết nối có giúp ích giải quyết vấn đề này.

Từ giờ, tác giả triển khai thí nghiệm trên các nhóm user có các mức thưa khác nhau. Cụ thể, dựa trên số lượng tương tác trên 1 user, tác giả chia tập test làm 4 nhóm, mỗi nhóm có tổng các tương tác bằng nhau. Lấy tập Gowalla làm ví dụ, số lượng tương tác trên 1 user nhỏ hơn 24, 50, 117 và 1014. Hình 4 minh họa kết quả theo điểm ndcg@20 trong các nhóm user khác nhau của Gowalla, Yelp2018\* và Amazon-Book. Ta thấy rằng điểm recall@20 có 1 xu hướng chung, bỏ qua tác động của giới hạn không gian. Ta có:

A picture containing text, line, screenshot, diagram

Description automatically generated

* NGCF và HOP-Rec đều có hiệu năng vượt trội so với tất cả các baseline trên tất cả các nhóm user. Nó chứng tỏ rằng việc phát hiện được các kết nối bậc cao rất có ảnh hưởng đến việc học biểu diễn cho các inactive user, khi mà các thông tin tương tác được phát hiện 1 cách hiệu quả. Vì vậy, nó rất hứa hẹn trong việc giải quyết vấn đề thưa trong hệ khuyến nghị.
* Phân tích cùng lúc các hình 4a, 4b, 4c, ta thấy rằng có sự cải thiện hiệu năng đáng kể trong 2 nhóm đầu (cao hơn 8.49% và 7.79% so với baseline tốt nhất với các nhóm user <24 và <50 trong tập Gowalla). Nhưng trong các nhóm khác thì hiệu năng cải thiện ít hơn (ví dụ với nhóm <1014 của Gowalla, hiệu năng chỉ tăng 1.29%). Chứng tỏ rằng, lan truyền embedding sẽ thể hiện tác động khác nhau trên các nhóm user khác nhau.

## **4.4. Nghiên cứu về NGCF (RQ2)**

Vì lớp lan truyền embedding đóng vai trò then chốt trong NGCF, ta sẽ nghiên cứu tác động của nó lên hiệu năng. Ta bắt đầu bằng việc xem xét tác động của số lượng lớp. Sau đó, ta sẽ nghiên cứu cách thức mà ma trận Laplace tác động lên hiệu năng ( ví dụ hệ số pui giữa user u và item i). Ngoài ra, ta cũng sẽ phân tích tác động của các hệ số chính như tỷ lệ node dropout, message dropout. Ta cúng sẽ nghiên cứu quá trình training của NGCF.

### **4.4.1. Tác động của số lượng lớp**

Để nghiên cứu vấn đề này, ta dùng các độ sâu lớp khác nhau. Cụ thể, ta tìm kiếm số lớp trong tập {1, 2, 3, 4}. Bảng 3 tổng hợp các kết quả thực nghiệm, trong đó NGCF-n là mô hình có n lớp lan truyền embedding. Kết hợp thông tin ở bảng 2 và bảng 3, ta thấy:

A picture containing text, screenshot, font, number

Description automatically generated

* Khi tăng độ sâu của NGCF sẽ cải thiện kết quả khuyến nghị. Rõ ràng là NGCF-2 và NGCF-3 có kết quả tốt hơn NGCF-1 (chỉ xem xét các hàng xóm bậc 1) trên tất cả các trường hợp. Ta cho rằng sự cải thiện này là do hiệu quả của việc mô hình hóa hiệu ứng CF: sự tương đồng của user tương tác và thông tin tương tác được chứa trong lần lượt kết nối bậc 2 và bậc 3.
* Khi xếp chông thêm lớp lan truyền sau lớp NGCF-3 ta thấy rằng NGCF-4 có hiện hộng overfitting trên tập Yelp2018\*. Điều này có tể là do khi triển khai kiến trúc quá sâu đã làm nhiễu ảnh hưởng đến việc học biểu diễn. Với 2 tập dữ liệu còn lại cũng không cải thiện được thêm kết quả khi có lớp thứ 4, cho thấy việc triển khai 3 lớp là phù hợp để bắt được thông tin CF.
* Khi sử dụng số lượng lớp lan truyền khác nhau, NGCF đều cho kết quả tốt hơn so với các phương pháp khác trên cả 3 tập dữ liệu. Điều này chứng minh tính hiệu quả của NGCF, cho thấy mô hình hóa tường minh các kết nối bậc cao có thể hỗ trợ nhiệm vụ khuyến nghị.

### **4.4.2. Tác động của lớp lan truyền embedding và cơ chế tổng hợp lớp**

Để nghiên cứu cách thức lớp lan truyền embedding (ví dụ như tích chập đồ thị) tác động lên hiệu năng, ta thử đặt NGCF-1 trên các lớp khác nhau. Cụ thể, ta loại bỏ tương tác biểu diễn giữa 1 node và hàng xóm của nó trong message passing function (công thức 3) và đặt nó như trong PinSage và GC-MC (gọi tên tương ứng là NGCF-1PinSage và NGCF-1GC-MC. Ngoài ra, theo SVD++, ta cũng có thể đặt tương tác này theo công thức (12), gọi là NGCF-1SVD++. Kết quả được trình bày trong bảng 4, ta thấy rằng:

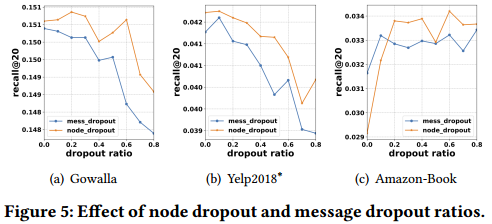
A picture containing text, screenshot, font, number

Description automatically generated

* NGCF-1 là tốt nhất so với các biến thể khác. Ta cho rằng sự cải thiện này nằm ở tương tác biểu diễn làm cho các message được lan truyền phụ thuộc vào sự tương tự giữa ei, eu và cơ chế thực hiện (ví dụ cơ chế attention [2]. Trong đó, tất cả các biến thể chỉ xem xét các biến đổi tuyến tính. Điều này chứng tỏ sự phù hợp và tính hiệu quả của phương pháp lan truyền embedding tác giả đề xuất.
* Trong hầu hết trường hợp, NGCF-1SVD++ có hiệu năng kém hơn NGCF-1PinSage và NGCF-1GC-MC. Nó minh họa tính quan trọng của message đi qua các node và của biến đổi phi tuyến.
* Kết hợp thông tin tại bảng 2 và bảng 4, ta thấy rằng khi ghép tất cả các đầu ra của các lớp với nhau, NGCF-1PinSage và NGCF-1GC-MC đạt được hiệu năng tốt hơn so với PinSage và GC-MC. Điều này nhấn mạnh tầm quan trọng của cơ chế tổng hợp lớp.

### **4.4.3. Tác động của Dropout**

Dựa vào các kết quả nghiên cứu trước [29], tác giả triển khai node dropout và message dropout để tránh cho NGCF khỏi overfitting. Hình 5 biểu diễn tác động của tỷ lệ message dropout p1 và tỷ lệ node dropout p2 trên nhiều độ đo đánh giá và nhiều tập dữ liệu.



Giữa 2 chiến thuật dropout, node dropout cho hiệu năng tốt hơn. Ví dụ với tập dữ liệu Gowalla, đặt p2=0.2 cho điểm recall@20 cao nhất là 0.1514, tốt hơn điểm cao nhất của message dropout là 0.1506. Lý do có thể là khi loại bỏ tất cả các message đầu ra của 1 user hay item cụ thể làm cho biểu diễn ít bị ảnh hưởng bởi các cạnh và node nào đó. Vì vậy, node dropout hiệu quả hơn message dropout, vốn phù hợp với việc tìm ra prior effort [29]. Đây là 1 phát hiện thú vị cho thấy node dropout có thể là chiến lược hiệu quả giải quyết vấn đề overfitting của mạng đồ thị.

### **4.4.4. Kiểm tra hiệu năng trên epoch**

Hình 6 biểu diễn kết quả kiểm tra hiệu năng với điểm recall tại mỗi epoch của phương pháp MF và NGCF. Vì giới hạn không gian, tác giả bỏ qua việc kiểm tra hiệu năng bằng điểm ndcg với kịch bản tương tự. Có thể thấy rằng, NGCF có tốc độ hội tụ nhanh hơn MF trên cả 3 tập dữ liệu. Điều này là hợp lý vì các user và item không kết nối trực tiếp đều được đề cập trong quá trình tối ưu các cặp tương tác trong mỗi mini batch. Quan sát này chứng minh khả năng mô hình hóa tốt hơn của NGCF và tính hiệu quả của việc triển khai lan truyền embedding trong không gian embedding.

A graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of a graph of

Description automatically generated with low confidence

## **4.5. Hiệu quả của kết nối bậc cao (RQ3)**

Phần này, ta cố gắng giải thích vì sao lớp lan truyền embedding tạo điều kiện việc học biểu diễn trong không gian embedding. Từ đây, ta chọn ngẫu nhiên 6 user trong tập Gowalla cùng với các item liên quan. Ta quan sát các biểu diễn của chúng bị ảnh hưởng bởi chiều sâu của NGCF.

A picture containing text, screenshot

Description automatically generated

Hình 7a, 7b cho thấy hình ảnh trực quan của các biểu diễn từ MF (NGCF-0) và NGCF-3. Lưu ý rằng các item này đều thuộc tập test, nghĩa là không tương ứng với các user trong pha training. Có 2 điểm mấu chốt ở đây là:

* Kết nối giữa các user và item được thể hiện tốt trong không gian embedding, nghĩa là chúng gần nhau trong không gian. Cụ thể, biểu diễn của NGCF-3 trình bày việc phân cụm rất rõ ràng, các điểm cùng màu (chính là các item được mua bởi cùng user) được nhóm vào 1 cụm.
* Xem xét cùng 1 user (ví dụ 12201 và 6880) trong hình 7a và 7b, ta thấy rằng khi xếp chồng 3 lớp lan truyền embedding, kết quả embedding của các item cũ sẽ gần nhau hơn. Điều này chứng minh lớp lan truyền embedding được tác giả đề xuất có thể mang các thông tin tương tác ẩn (qua NGCF-3) đưa vào kết quả biểu diễn.

# **5. Kết luận và hướng phát triển**

Trong nghiên cứu này, tác giả đã đưa thông tin tương tác tường minh vào quá trình embedding của model-based CF. Tác giả đề xuất 1 framework mới là NGF, có thể đạt ddwwocj mục tiêu bằng cách tận dụng các kết nối bậc cao trong đồ thị tích hợp user-item. Điểm mấu chót của NGCF là lớp lan truyền embedding, mà dựa vào đó cho phép sự tương tác giữa embedding của user và item để thu được thông tin tương tác. Các thí nghiệm trên 3 tập dữ liệu thật đã chứng minh sự phù hợp và hiệu quả của việc đưa cấu trúc đồ thị user-item vào quá trình embedding. Trong tương lai, tác giả muốn cải thiện MGCF bằng cách thêm vào cơ chế attention [2] để học các trọng số hàng xóm 1 cách trọng tâm cho quá trình lan truyền embedding và cho các kết nối tại các bậc khác nhau. Việc này rất có lợi trong việc sinh và diễn giải mô hình. Hơn nữa, tác giả cũng quan tâm đến việc học đối nghịch (adversarial learning) [13] trong quá trình embedding user/item, và cấu trúc đồ thị để tăng cường tính bền vững của NGCF.

Nghiên cứu này trình bày khởi đầu 1 hướng mới trong việc khám phá các thông tin có cấu trúc với cơ chế message-passing trong model-based CF và mở ra 1 khả năng nghiên cứu mới. Đặc biệt, có khá nhiều dạng thông tin có cấu trúc khác sẽ hữu ích cho việc hiểu các hành vi của user, như các thuộc tính chéo [41] trong hệ khuyến nghị với ngữ cảnh nhận thức và giàu ngữ nghĩa (context-aware and semantics-rich recommendation [22, 27], đồ thị thông tin item [32] và mạng xã hội [34]. Ví dụ, bằng việc tích hợp đồ thị thông tin item với đồ thị user-item, ta cso thể thiết lập kết nối thông tin giữa user và item, từ đó giúp cho quá trình ra quyết định của user trong việc lựa chọn item.