

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



Cù Nguyên Giáp

**NGHIÊN CỨU PHÁT TRIỂN HỆ TƯ VẤN NHÓM
THEO TIẾP CẬN MỜ TRỰC CẢM
VÀ TÍCH PHÂN CHOQUET**

TÓM TẮT LUẬN ÁN TIẾN SĨ MÁY TÍNH

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Hà Nội - 2025

**Công trình được hoàn thành tại: Học viện Khoa học và Công nghệ,
Viện Hàn lâm Khoa học và Công nghệ Việt Nam**

Người hướng dẫn khoa học:

Người hướng dẫn 1: PGS.TS. Lê Hoàng Sơn, Viện Công nghệ thông tin, Đại học Quốc gia Hà nội

Người hướng dẫn 2: TS. Nguyễn Như Sơn, Viện Công nghệ thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

Phản biện 1: PGS.TS. Phạm Văn Hải

Phản biện 2: PGS.TS. Hoàng Việt Long

Phản biện 3: PGS.TS. Trần Giang Sơn

Luận án được bảo vệ trước Hội đồng đánh giá luận án tiến sĩ cấp Học viện họp tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam vào hồi 9 giờ 00, ngày 11 tháng 2 năm 2025.

Có thể tìm hiểu luận án tại:

1. Thư viện Học viện Khoa học và Công nghệ
2. Thư viện Quốc gia Việt Nam

MỞ ĐẦU

1. Tính cấp thiết của nghiên cứu

Trong những năm gần đây, Hệ tư vấn (HTV) được coi là giải pháp lọc thông tin cho người dùng khi các hệ thống có lượng dữ liệu quá lớn và cách tìm kiếm dữ liệu theo từ khóa không thực sự phù hợp. Thực tế cho thấy HTV đã được nghiên cứu, phát triển và ứng dụng trong hầu hết các HTTT có số lượng người dùng lớn hiện nay [1]. Trong đó, hệ tư vấn hướng tới giải quyết bài toán đưa ra các khuyến nghị phù hợp cho một nhóm người dùng được gọi là Hệ tư vấn nhóm (HTVN) [2]. Về mặt mô hình hóa, HTVN là một mô hình khái quát của HTV đơn người dùng, HTVN sẽ trở thành hệ tư vấn đơn người dùng khi mỗi nhóm chỉ có duy nhất một thành viên.

Hệ tư vấn nhóm ngày càng trở thành một lĩnh vực nghiên cứu quan trọng, đánh dấu từ những nghiên cứu và ứng dụng đầu tiên của Masthoff năm 2004 [3], [4] về ứng dụng HTVN vào tư vấn lựa chọn chương trình truyền hình, và các nghiên cứu ứng dụng vào các lĩnh vực như du lịch, dịch vụ giải trí [2], [5]-[8]. HTVN sẽ ngày càng trở nên phổ biến hơn khi nhu cầu ra quyết định chung cho nhóm người dùng trong các hoạt động cộng tác trở nên phổ biến [9].

Các nghiên cứu về hệ tư vấn nhóm có thể chia làm hai nhóm tiếp cận chính: (1) Tiếp cận hợp người dùng và (2) Tiếp cận hợp khuyến nghị. Tổng quan tài liệu cho thấy cách tiếp cận thứ hai chiếm ưu thế rất lớn so với cách tiếp cận thứ nhất.

Có nhiều tiêu chí được sử dụng để đánh giá một HTVN [16], như độ chính xác trong dự báo, sự đa dạng, độ bao phủ hay sự đồng thuận và tính công bằng. Trong nghiên cứu về hệ tư vấn nhóm cho thấy khi nghiên cứu về HTVN xu hướng ưu tiên *tính công bằng* của khuyến nghị là rất quan trọng.

Bên cạnh đó, để xây dựng hệ tư vấn nhóm sát với thực tiễn, có thể thấy nghiên cứu về *hệ tư vấn nhóm động theo tiếp cận tính toán mờ* cần được nghiên cứu sâu rộng hơn. Kết hợp hai yếu tố “động” và “tính toán mờ” có thể giúp bài toán HTVN biểu diễn đúng đặc trưng về sự bất định và tính không chắc chắn khi đưa ra các đánh giá của người dùng, và sự biến động trong đánh giá của người dùng, sự thay đổi về tính hấp dẫn của sản phẩm theo thời gian, từ đó giúp mô hình gần với thực tiễn hơn.

Dựa trên các công bố về HTVN, HTVN theo tiếp cận tính toán mờ, và HTVN động hiện nay thì vẫn còn tồn tại một số hạn chế nhất định. Do đó, trong luận án này, NCS đề xuất thực hiện “***Nghiên cứu phát triển hệ tư vấn nhóm theo tiếp cận mờ trực cảm và tích phân Choquet***”. Trong đó sẽ phát triển mô hình hệ tư vấn cho nhóm người dùng, trong đó sử dụng độ đo mờ để cải tiến độ công bằng trong khuyến nghị cho nhóm người dùng và áp dụng lý thuyết tập mờ mở rộng để biểu diễn và xử lý tốt hơn thông tin có tính do dự, bất định trong phản hồi và đánh giá của người dùng, đồng thời xem xét tính động trong hệ tư vấn nhóm.

2. Mục tiêu nghiên cứu của luận án

Mục tiêu chung: Nghiên cứu phát triển hệ tư vấn nhóm động theo tiếp cận mờ trực cảm đảm bảo sự công bằng trong khuyến nghị.

3. Nội dung nghiên cứu chính của luận án

Nội dung chính của luận án gồm ba phần tương đương với ba chương nội dung. Trong đó: **Chương 1** trình bày các khái niệm cơ bản trong lý thuyết về hệ tư vấn nhóm và các vấn đề liên quan. Trên cơ sở đó, luận án phân tích các vấn đề còn tồn tại và nêu rõ các mục tiêu nghiên cứu, cách thức giải quyết và các kết quả đạt được của luận án

một cách ngắn gọn. **Chương 2** trình bày nghiên cứu về hệ tư vấn nhóm với tiếp cận xem xét tới sự công bằng dựa trên một độ đo mờ. Kết hợp giữa hai mục tiêu về tổng lợi ích của các thành viên và tính công bằng giữa các thành viên chúng ta sẽ phải giải quyết bài toán tối ưu đa mục tiêu trong HTVN. **Chương 3** trình bày đề xuất sử dụng lý thuyết mờ trực cảm để phát triển một hệ tư vấn nhóm động. Từ đó, một đề xuất về hệ tư vấn nhóm động dựa trên tập mờ trực cảm phát triển, đồng thời trong hệ tư vấn này một phép toán hợp với tích phân Choquet cho tập mờ trực cảm tiếp tục được đề xuất và thử nghiệm để tìm được một mô hình HTVN phù hợp với thực tiễn, và có khả năng ứng dụng cao.

Chương 1. TỔNG QUAN VỀ HỆ TƯ VẤN NHÓM

1.1. Giới thiệu về Hệ tư vấn nhóm

1.1.1. Hệ tư vấn nhóm

Các Hệ tư vấn đầu tiên được tập trung phát triển để đưa ra các khuyến nghị cho cá nhân, tuy nhiên, hiện nay các hệ thống khuyến nghị còn hướng đến khuyến nghị cho một nhóm người dùng. Ứng dụng của Hệ tư vấn nhóm ngày càng mở rộng theo thời gian [2], [3], [30], [31].

Khái niệm hệ tư vấn nhóm: Có thể hiểu Hệ tư vấn nhóm là hệ tư vấn đưa ra một tập các đối tượng (sản phẩm, dịch vụ .v.v.) được cho là phù hợp tới một nhóm người dùng [4]. Hệ tư vấn nhóm đơn giản nhất có thể mô hình hóa như sau.

Mô hình hóa Hệ tư vấn nhóm

Cho $U = \{u_1, u_2, \dots, u_n\}$ và $I = \{i_1, i_2, \dots, i_m\}$ là tập hợp người dùng và sản phẩm, dịch vụ; cho $R \subseteq U \times I \rightarrow D$ là tập hợp các đánh giá của người dùng cho các sản phẩm trên miền đánh giá D . Cho nhóm người dùng $g = \{u_1, u_2, \dots, u_l\} \mid u_i \in U$, khi đó hệ khuyến nghị nhóm được mô hình như sau:

$$HTVN(g, I) = \arg \max_{i \in I} [pref(g, i)] \quad (1.1)$$

Trong đó $pref(g, i)$ là giá trị dự báo thể hiện đánh giá của nhóm người dùng g cho một sản phẩm, dịch vụ i .

1.1.2. Tổng quan các nghiên cứu về Hệ tư vấn nhóm

Hệ tư vấn nhóm có thể coi là bắt đầu được phát triển từ những năm cuối của thập niên 1990, đầu thập niên 2000, với nghiên cứu nổi bật của Mathoff và cộng sự [4], và sau đó đến những năm gần đây hệ tư vấn nhóm mới thực sự trở thành một nhánh nghiên cứu nổi bật.

Các nghiên cứu ban đầu về HTVN tập trung chủ yếu vào việc tổng hợp các sở thích cá nhân để tạo ra các đề xuất cho nhóm [4], [10]. Sau đó các kỹ thuật lọc cộng tác vào Hệ thống tư vấn nhóm [6], [11], HTVN có tích hợp mô hình hóa ảnh hưởng xã hội [12]-[14], và HTVN đặt trọng tâm là nâng cao sự đa dạng và tính công bằng trong khuyến nghị [17], [18], [33] dần được phát triển.

Bên cạnh đó, việc cải thiện cách thức giải quyết bài toán độ công bằng giữa những người dùng trong nhóm người dùng sẽ làm tăng sự thỏa mãn nói chung của người dùng, từ đó làm tăng khả năng ứng dụng thực tiễn của các Hệ tư vấn nhóm [21].

Có hai cách tiếp cận phổ biến trong HTVN, đó là cách tiếp cận “hợp hồ sơ người dùng” và cách tiếp cận “hợp khuyến nghị”. Hướng tiếp cận thứ hai, và cũng là hướng tiếp cận phổ biến hơn trong HTVN hiện nay [14]. Trong cách tiếp cận này, tính công bằng của việc tạo kiến nghị được quyết định trong giai đoạn “đồng thuận” của hệ thống khuyến nghị.

1.1.3. Hệ tư vấn nhóm sử dụng cách tiếp cận hợp khuyến nghị đơn của người dùng trong nhóm

Trong pha đồng thuận, các toán tử hợp khác nhau được sử dụng phản ánh các chiến lược trong việc xây dựng một giá trị thể hiện sự

yêu thích của cả nhóm đối với một sản phẩm dựa trên đánh giá của từng cá nhân trong nhóm. Các tiếp cận chính được các nghiên cứu trước đây sử dụng trong pha này bao gồm “chiến lược tối đa hóa tổng lợi ích”, ví dụ như “chiến lược cộng”, “chiến lược trung bình” và “chiến lược nhân”; chiến lược dựa trên người yếu thế hoặc người có ưu thế (“chiến lược ít thiệt thòi nhất” và “chiến lược ưu thế nhất”); hay các cơ chế dựa trên cách thức biểu quyết trong thực tế như “chiến lược biểu quyết” và “chiến lược luật Copeland” hay một chiến lược cân bằng hơn như “chiến lược chỉ số Borda” và “chiến lược công bằng” [2], [32].

1.1.4. Các phương pháp đánh giá Hệ tư vấn nhóm.

Hiệu quả của Hệ tư vấn nhóm có thể được đánh giá thông qua các độ đo phản ánh một trong các khía cạnh sau: Độ chính xác phân lớp, độ chính xác dự báo, độ chính xác trong phân hạng, độ bao phủ và tính ngẫu nhiên, độ đồng thuận và công bằng [16]. Trong đó, độ đo tính đồng thuận và độ công bằng ngày càng được xét đến nhiều hơn trong Hệ tư vấn nhóm.

Khái niệm công bằng trong Hệ tư vấn nói chung có thể đề cập đến sự công bằng giữa người dùng, sự công bằng giữa người cung cấp hay cả hai đối tượng trên [17], [18], [33]. Trong Hệ tư vấn nhóm, các nghiên cứu về tính công bằng nghiêng về khía cạnh khác biệt về mức độ thỏa mãn, hay đánh giá giữa các người dùng trong nhóm về các sản phẩm, dịch vụ được khuyến nghị. Một số nghiên cứu gần đây đã đề xuất những định nghĩa và độ đo cho khái niệm công bằng trong Hệ tư vấn nhóm, nhưng các nghiên cứu sâu về lĩnh vực này hiện còn thiếu.

1.2. Tổng quan nghiên cứu về HTVN động và mờ.

1.2.1. Hệ tư vấn nhóm Động

Trong Hệ tư vấn nói chung, các phương pháp khai thác thông tin xem xét đến sự biến động theo thời gian có thể được tạm chia thành bốn loại [42]. Mỗi loại thể hiện một quan điểm khác nhau khi khai thác thông tin động và theo thời gian. Bốn cách tiếp cận gồm: 1) Tiếp cận gần đúng; 2) Tiếp cận dựa trên cơ sở gom nhóm; 3) Phương pháp cập nhật trực tuyến và 4) Tiếp cận dựa trên tính động. Trong đó, cách tiếp cận dựa trên tính động được ứng dụng phổ biến. Cách tiếp cận này dựa trên mô hình hóa rõ ràng các biến động theo thời gian trong phản hồi để theo dõi xu hướng thay đổi của các yếu tố như sở thích của người dùng và mức độ hấp dẫn của sản phẩm, dịch vụ [43, 44].

Tổng quan nghiên cứu cho thấy HTVN là một vấn đề được nghiên cứu sau và các nghiên cứu về HTVN trước đây thường tập trung vào giải quyết bài toán kết hợp các đánh giá của thành viên để tạo ra đánh giá của nhóm. Các nghiên cứu về HTVN có sử dụng tiếp cận thông tin động còn tương đối hạn chế. Một số nghiên cứu điển hình có thể chỉ ra như nghiên cứu của Jinpeng Chen và cộng sự [52], hay nghiên cứu của Huang về pha đồng thuận của HTVN có xem xét mối quan hệ thứ bậc của các sản phẩm theo thời gian [53]. Có thể thấy nghiên cứu HTVN theo tiếp cận thông tin động sẽ phản ánh tốt hơn thực tế các thông tin trong hệ thống.

1.2.2. Hệ tư vấn nhóm theo tiếp cận tính toán mờ

Sử dụng tiếp cận tính toán mờ trong xây dựng hệ tư vấn là một chiến lược đã được nghiên cứu rộng rãi. Tiếp cận này có nhiều ưu điểm như là có thể biểu diễn và xử lý thông tin không chắc chắn được thể hiện trong đánh giá của người dùng với sản phẩm [15], [35], [54]. Nghiên cứu ứng dụng trực tiếp tính toán mờ trong hệ tư vấn nhóm có

phần hạn chế hơn so với ứng dụng tính toán mờ trong hệ tư vấn đơn người dùng.

Tổng quan nghiên cứu cho thấy phát triển HTVN theo tiếp cận động và phát triển HTVN theo tiếp cận tính toán mờ đều có những ưu điểm nổi bật, và có thể hỗ trợ lẫn nhau. Tuy nhiên, các nghiên cứu về các tiếp cận này còn khá thiếu và cần được tìm hiểu sâu hơn, để từ đó xây dựng được một mô hình HTVN tốt hơn.

1.2.3. Một số khái niệm trong lý thuyết tập mờ trực cảm

1.2.3.1. Khái quát về Tập mờ trực cảm

Trong các tập mờ và tập mờ được nghiên cứu, tập mờ trực cảm có những ưu thế nhất định trong biểu diễn và xây dựng các hệ tư vấn [55]. Định nghĩa về tập mờ trực cảm được giới thiệu bởi Atanasov [56], [57].

Định nghĩa 1.1: cho không gian nền X , một tập mờ trực cảm A trên X được cho bởi:

$$A = \{(x, \mu_A(x), \nu_A(x)) \mid x \in X\} \quad (1.17)$$

Thỏa mãn: $\mu_A : X \rightarrow [0,1], \nu_A : X \rightarrow [0,1]$. Trong đó, $\mu_A(x) \in [0,1]$ thể hiện độ thuộc của một phần tử x với A , và $\nu_A(x) \in [0,1]$ thể hiện độ không thuộc của x với A , và điều kiện $0 \leq \mu_A(x) + \nu_A(x) \leq 1$ thỏa mãn $\forall x \in X$.

Các phép toán đại số cho tập mờ trực cảm đã được giới thiệu trong [58]. Các phép toán đại số trong tập mờ trực cảm ở trên là nền tảng để phát triển các thuật toán xử lý với dữ liệu mờ trực cảm.

1.2.3.2. Khoảng cách và độ tương tự cho tập mờ trực cảm

Khoảng cách Hamming: Cho A, B là hai tập mờ trực cảm trên tập nền $X = \{x_1, \dots, x_n\}$

$$d_{IFS}(A, B) = \sum_{i=1}^n (|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\pi_A(x_i) - \pi_B(x_i)|) \quad (1.20)$$

Khoảng cách Euclidean:

$$e_{IFS}(A, B) = \sqrt[2]{\sum_{i=1}^n ((\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2 + (\pi_A(x_i) - \pi_B(x_i))^2)} \quad (1.21)$$

1.2.3.3. Phép toán trung bình với các số mờ trực cảm

Dưới đây là các phép toán hợp và trung bình quan trọng.

Cho $A = \{(\mu_{a_i}, \nu_{a_i}); i = 1, 2, \dots, n\}$ là tập các số mờ trực cảm. Ta có:

Trung bình số học có trọng số:

$$IFAW(A, W) = \bigoplus_{i=1}^n a_i w_i = \left\langle 1 - \prod_{i=1}^n (1 - \mu_{a_i})^{w_i}, \prod_{i=1}^n \nu_{a_i}^{w_i} \right\rangle \quad (1.30)$$

Trung bình Bonfferroni (Intuitionistic Bonfferroni Mean) [62]

$$IFB^{p,q}(A) = \left(\frac{1}{n(n-1)} \left(\bigoplus_{\substack{i,j=1 \\ i \neq j}}^n a_i^p \otimes a_j^q \right) \right)^{\frac{1}{p+q}} \quad (1.31)$$

1.2.4. Tích phân Choquet

Để xây dựng được phép toán hợp cho pha đồng thuận của hệ tư vấn nhóm dựa trên tích phân Choquet, chúng ta cần trình bày lại cách thức tính tích phân Choquet.

a) Tích phân Choquet:

Định nghĩa 2.2: tích phân Choquet của một vector $r \in \mathbb{R}^m$ theo một hàm dung lượng ξ được định nghĩa như sau:

$$CQ_{\xi}(r) = \sum_{i=1}^m (r_i^{\uparrow} - r_{i-1}^{\uparrow}) \xi(Y_i^{\uparrow}) \quad (1.33)$$

Trong đó $r^\uparrow = \{r_1^\uparrow, \dots, r_m^\uparrow\}$ là một hoán vị theo thứ tự tăng dần của r trong đó $0 = r_0^\uparrow \leq r_1^\uparrow \leq \dots \leq r_m^\uparrow$, và tập hợp $Y_i^\uparrow = \{j \in M, r_j \geq r_i^\uparrow\} = \{i^\uparrow, (i+1)^\uparrow, \dots, m^\uparrow\}$ với $i < m$ và $Y_{m+1}^\uparrow = 0$.

b) Độ đo mờ

Khi xây dựng một phép toán hợp dựa trên tích phân Choquet, nếu hàm dung lượng là một hàm phi cộng tính thì phép hợp này thể hiện một độ đo mờ phản ánh mục tiêu của phép toán hợp [41,64]. Xây dựng một độ đo mờ tối ưu là bài toán thuộc lớp NP-complete do đó không khả thi để giải với các thuật toán chính xác.

Trong Hệ tư vấn nhóm, để áp dụng phép toán hợp dựa trên tích phân Choquet, nó đòi hỏi phải tính toán giá trị hàm dung lượng cho từng nhóm người dùng trong thời gian phù hợp. Cách tiếp cận khả thi là đề xuất một thuật toán đánh giá trực tiếp giá trị của các điểm trong phần mở rộng tuyến tính cần thiết và nó phải là một quy trình tính toán hiệu quả. Luận án này thực hiện theo cách tiếp cận này và xác định một số hàm dung lượng dựa trên sự tương tác của người dùng và mục tiêu gia tăng tính công bằng trong các khuyến nghị của HTVN.

1.3. Kết luận

Chương 1 trình bày một số khái niệm nền tảng về hệ tư vấn nhóm dựa trên cách thức khái quát hóa hệ tư vấn đơn người dùng. Trong tổng quan nghiên cứu về HTVN bao gồm các chiến lược tiếp cận, phương thức đánh giá, các nghiên cứu sử dụng thông tin dạng tĩnh, thông tin dạng động và các nghiên cứu sử dụng tiếp cận tính toán mờ được trình bày. Các tiếp cận được trình bày và phân tích theo các ưu nhược điểm của mỗi phương pháp. Trên cơ sở đó, chương 1 đưa ra các vấn đề nghiên cứu của luận án. Cụ thể luận án tập trung nghiên cứu về HTVN và đề xuất, phát triển thuật toán HTVN sử dụng độ đo mờ dựa trên tích phân Choquet nhằm cải thiện độ công bằng của

khuyến nghị trong chương 2 và đề xuất, phát triển HTVN mờ động trong chương 3.

Chương 2. TĂNG CƯỜNG SỰ CÔNG BẰNG TRONG HỆ TƯ VẤN NHÓM VỚI ĐỘ ĐO MỜ

2.1. Mở đầu

Với HTVN, vấn đề công bằng trong khuyến nghị là một vấn đề được đặc biệt quan tâm [16], [18], [66]. Các nghiên cứu này đưa ra một số đề xuất, trong đó coi sự công bằng của HTVN là tỷ lệ số người hài lòng trên tổng số thành viên nhóm [10], độ lệch của mức độ hài lòng của các thành viên trong nhóm [19] hoặc coi công bằng của tập sản phẩm được đề xuất là một “gói” hơn là hơn một tập hợp các sản phẩm độc lập [18].

Bên cạnh đó, một thách thức khác được đặt ra trong việc tìm kiếm giải pháp công bằng tốt trong HTVN dựa trên sự đồng thuận là sự ưa thích của thành viên đối với một sản phẩm, dịch vụ bị ảnh hưởng bởi sự tương tác của các thành viên [3], [67], [68]. Do đó, để ước tính sự mất cân bằng giữa sở thích của các thành viên trong nhóm cần tính đến ương tác của các thành viên.

Trong pha đồng thuận, thay thế cho các phép toán hợp đã có trước đó, luận án đề xuất sử dụng tích phân Choquet để tạo đề xuất nhóm trong giai đoạn đồng thuận của HTVN. Phép toán hợp dựa trên tích phân Choquet mở rộng phạm vi tìm kiếm giải pháp so với các phép hợp có trọng số và nó có thể đưa ra khuyến nghị cân bằng hơn so với các chiến lược trước đó nhờ xây dựng một độ đo mờ phù hợp [41], [70].

2.2. Đề xuất mô hình HTVN dùng toán tử hợp dựa trên tích phân Choquet

2.2.1. Mô hình HTVN với toán tử hợp dựa trên tích phân Choquet

a) Đề xuất mô hình HTVN dựa trên tích phân Choquet

Trong luận án này, NCS đề xuất mô hình HTVN với hai pha, pha thứ nhất là pha sinh khuyến nghị sẽ dự báo đánh giá của một người dùng với các sản phẩm, và pha thứ 2 là pha thể hiện cơ chế đồng thuận giữa các thành viên trong một nhóm người dùng. Cụ thể:

- Pha khuyến nghị (recommendation phase) sử dụng lọc cộng tác theo người dùng.
- Pha đồng thuận: sử dụng toán tử hợp dựa trên tích phân Choquet để ước lượng đánh giá của nhóm cho sản phẩm, dịch vụ dựa trên đánh giá của từng thành viên. Dựa trên kết quả đánh giá của nhóm sẽ đưa ra khuyến nghị cho nhóm theo nguyên tắc lựa chọn N sản phẩm được đánh giá cao nhất.

b) Các hàm dung lượng được đề xuất

Hàm dung lượng thứ nhất:

Dưới đây, các tác giả đề xuất một hàm dung lượng được gọi là “hàm dung lượng thứ nhất” dựa trên mức độ tương tác của người dùng với hệ thống. Nghiên cứu này có dựa trên đề xuất trong nghiên cứu của tác giả Huynh và cộng sự [78].

Cho nhóm người dùng g được xem như là điều kiện khi lựa chọn các sản phẩm. Hàm dung lượng được định nghĩa như sau:

$$\xi(u_i) = \begin{cases} \frac{\omega(u_i)}{\sum_{u_i \in g} \omega(u_i)}, & \text{if } \sum_{u_i \in g} \omega(u_i) < 1 \\ \omega(u_i), & \text{otherwise} \end{cases} \quad (2.3)$$

Trong đó:

$$\omega(u_i) = \frac{\text{count}(r_{u_i j} \geq 0)}{|I|}, j = 1, \dots, |I| \quad (2.4)$$

và giá trị hàm dung lượng của một tập con $A \subset g$ là:

$$\xi(A) = \sum_{u_i \in A} \xi(u_i) + \sigma(A) \quad (2.5)$$

Nếu $\xi(A) > 1$ nó được gán bằng 1.

Định lý 2.1: Hàm dung lượng được định nghĩa ở trên thỏa mãn tính chất cộng tính khi $\sigma(A) = 0$ và $\sum_{u_j \in g} \omega(u_j) \leq 1$.

Tính chất 2.1: Trong trường hợp hàm dung lượng được định nghĩa ở trên có tính chất cộng tính thì toán tử hợp dựa trên tích phân Choquet trở thành một phép toán hợp có trọng số (weighted sum aggregation).

Trong tiếp cận của hàm dung lượng thứ nhất ta có $\sigma(A) \neq 0$, và được tính như sau.

$$\sigma(A) = \sum_{u_i \in A} \omega'(u_i) \quad (2.6)$$

Trong đó:

$$\omega'(u_i) = \frac{\text{count}(r_{u_i j} > \text{aveRate}) - \text{count}(r_{u_i j} < \text{aveRate})}{|I|} \quad (2.7)$$

Tập sản phẩm I chứa tất cả các sản phẩm được đánh giá bởi u_i . $\omega'(u_i)$ thể hiện mức độ hứng thú của một người dùng trong việc đánh giá các sản phẩm.

Hàm dung lượng thứ 2

Trong các hoạt động nhóm, tương tác giữa các nhóm nhỏ người dùng trong đó sẽ ảnh hưởng tới sự thỏa mãn của một người. Với cùng một hoạt động, một người có thể cảm thấy thỏa mãn hơn khi tham gia

cùng với những người khác có tính tương đồng cao với mình. Do đó NCS đề xuất một hàm dung lượng mở rộng như sau.

$$\sigma'(A) = \frac{1}{dens(A)} * \sigma(A) \quad (2.8)$$

Mật độ (density) thể hiện sự tương đồng giữa một nhóm người dùng và nó dựa trên khoảng cách của mọi người dùng, và được tính dựa trên độ đo khoảng cách Mahalanobis.

2.2.2. Độ phức tạp thuật toán

Để so sánh độ phức tạp của thuật toán đề xuất mới cho pha đồng thuận và các chiến lược đồng thuận được trình bày trong mục 1.2.2, chúng ta cần đánh giá độ phức tạp của thuật toán dựa trên khái niệm O lớn. Thông qua đánh giá độ phức tạp từ giả mã của thuật toán, ta có kết luận về độ phức tạp của pha đồng thuận sử dụng toán tử hợp bằng tích phân Choquet sẽ là $O(|I|^2 \cdot |g| \cdot N)$.

2.3. Kết quả thực nghiệm và bàn luận

2.3.1. Dữ liệu thực nghiệm

Để so sánh phương pháp tiếp cận đề xuất và các chiến lược tiếp cận khác, dữ liệu MovieLens-1M được sử dụng. Do đó, khi ứng dụng trong bài toán HTVN cần xây dựng cơ chế sinh nhóm người dùng. Trong nghiên cứu này, với nhóm g_i một cơ chế lấy mẫu ngẫu nhiên được sử dụng để chọn $|g_i|$ người dùng (không lặp) để tạo các nhóm người dùng giả định.

2.3.2. Phương thức đánh giá

Giá trị thỏa mãn nhóm:

Chú ý rằng một mục tiêu của HTVN vẫn là tìm các sản phẩm có tổng độ thỏa mãn cao bên cạnh tính công bằng, do đó chúng ta cần độ

đo thể hiện giá trị thỏa mãn của cả nhóm người dùng. Công thức tính độ đo này như sau:

$$group_pref(g, I) = \frac{\sum_{i \in I} r_{g,i}}{|I|} \quad (2.13)$$

Trong đó $r_{g,i}$ thể hiện giá trị trung bình độ thỏa mãn của toàn bộ người dùng trong nhóm g cho một sản phẩm $i \in I$.

Độ đo tính công bằng: Để ước lượng tính công bằng của các khuyến nghị từ mô hình HTVN, luận án sử dụng 2 độ đo thể hiện trong công thức (1.14) như sau.

$$fairness(g, i) = \frac{\left| \bigcup_{u \in g} : r_{ui} \geq \theta \right|}{|g|}$$

Với θ là ngưỡng thể hiện người dùng hài lòng với sản phẩm.

Và độ đo thứ 2 thể hiện trong công thức

Error! Reference source not found. như sau:

$$fairness_{Var}(g, i) = 1 - Var\{r_{ui}, \forall u \in g\}$$

2.3.3. Kết quả và bàn luận

Một số kết quả chính được thể hiện như sau.

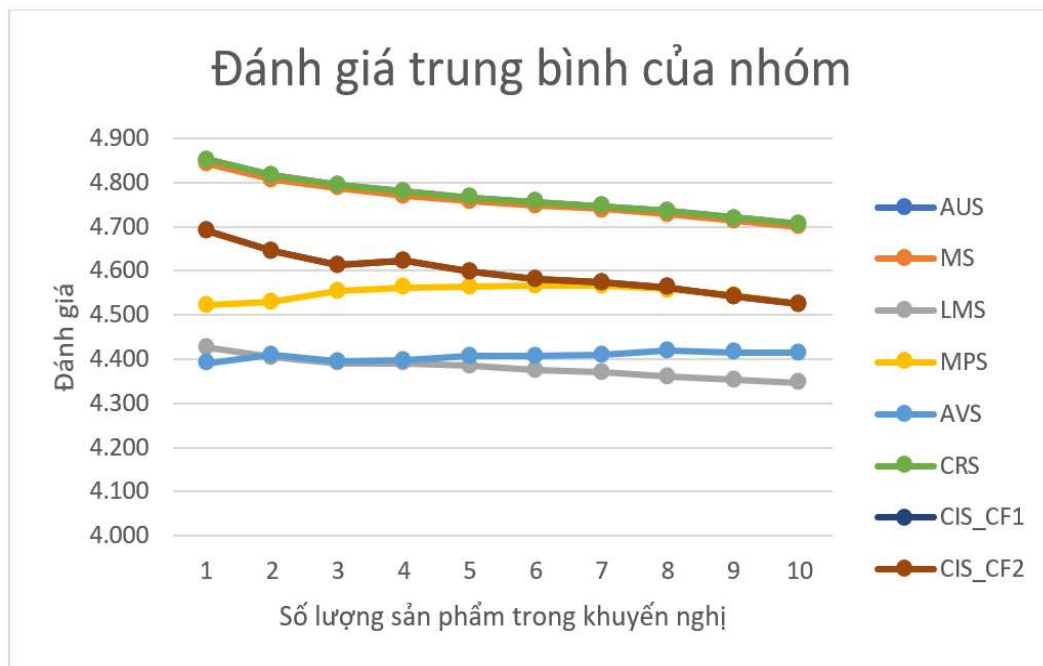
Bảng 2.4 Đánh giá trung bình của nhóm

Top-N	AUS	MS	LMS	MPS	AVS	CRS	CIS_CF1	CIS_CF2
N=1	4.717	4.694	4.373	4.384	4.412	4.717	4.556	4.560
N=2	4.689	4.670	4.350	4.404	4.383	4.689	4.536	4.537
N=3	4.669	4.658	4.339	4.387	4.357	4.669	4.514	4.516
N=4	4.652	4.645	4.337	4.396	4.350	4.652	4.503	4.505
N=5	4.637	4.628	4.331	4.397	4.349	4.637	4.481	4.483
N=6	4.625	4.615	4.327	4.401	4.342	4.625	4.465	4.467
N=7	4.614	4.605	4.313	4.402	4.341	4.614	4.451	4.453

N=8	4.600	4.594	4.301	4.398	4.339	4.600	4.442	4.444
N=9	4.586	4.578	4.292	4.390	4.335	4.586	4.433	4.433
N=10	4.572	4.564	4.279	4.378	4.331	4.572	4.420	4.420

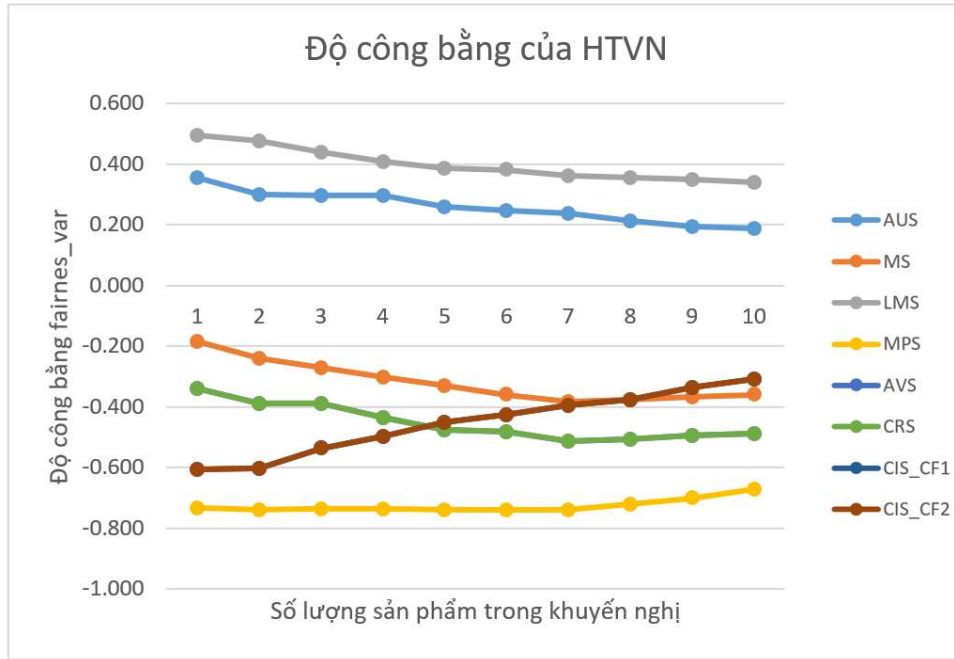
Bảng 2.4 cho thấy sự khác biệt của các mô hình về mức đánh giá trung bình của người dùng trong nhóm. Hiển nhiên, chiến lược AUS có kết quả tốt nhất khi xét đến mức độ thỏa mãn trung bình của toàn bộ người dùng trong nhóm. HTVN với chiến lược Copeland có cùng hiệu suất với AUS. Hai thuật toán đề xuất có kết quả rất gần với AUS và chúng tốt hơn các chiến lược khác.

Hình 2.2 thể hiện sự khác biệt về đánh giá trung bình của nhóm theo kích thước khuyến nghị cho thấy sự khác biệt của các thuật toán giảm dần khi kích thước khuyến nghị tăng.



Hình 2.2 Độ thỏa mãn trung bình của nhóm

Xét đến độ công bằng ta có độ công bằng của các chiến lược đồng thuận thể hiện trong hình sau:



Hình 2.5 Độ công bằng của HTVN

Chúng ta có thể thấy rằng “chiến lược ít thiệt thòi nhất” (LMS) và “chiến lược biểu quyết phê duyệt” (AVS) có yếu tố công bằng cao nhất, tuy nhiên hai chiến lược này ngược lại lựa chọn các sản phẩm mà sự hài lòng trung bình của người dùng là thấp nhất. “Chiến lược ưu thế” (MPS) đưa ra giải pháp công bằng thấp nhất trong mọi trường hợp và nó cũng đưa ra các khuyến nghị có mức độ ưa thích trung bình của nhóm rất thấp. “Chiến lược cộng ” (AUS) và chiến lược luật Copeland (CRS) đưa ra khuyến nghị có mức độ ưa thích trung bình nhóm người dùng cao nhất nhưng chúng cũng đưa ra khuyến nghị có tính công bằng thấp. HTVN với hai phép toán hợp dựa trên tích phân Choquet được đề xuất (CIS_CF1 và CIS_CF2 – tương ứng với hai hàm dung lượng khác nhau) không hơn các mô hình khác ở bất kỳ khía cạnh đơn lẻ nào trong hai khía cạnh: mức độ ưa thích trung bình của nhóm và hay độ đo công bằng. Mặc dù vậy, có thể dễ dàng nhận thấy rằng hai mô hình đề xuất này cân bằng giữa cả hai mục tiêu

trên. Do đó, mô hình HTVN được đề xuất sẽ có khả năng ứng dụng cao trong thực tế.

2.4. Kết luận chương 2

Chương 2 nghiên cứu xây dựng và phát triển hệ tư vấn nhóm dựa trên toán tử Choquet nhằm tăng sự công bằng của khuyến nghị. Mô hình đề xuất dựa trên tiếp cận xây dựng hệ tư vấn nhóm với pha đồng thuận. Trong pha đồng thuận, toán tử Choquet được sử dụng để phát triển một phép toán hợp các đánh giá đơn của người dùng trong nhóm. Với ứng dụng của toán tử Choquet, pha đồng thuận trong HTVN đã giải quyết được vấn đề tương tác giữa các thành viên trong nhóm, và biểu diễn được một độ đo mờ thể hiện trong phép toán hợp. Từ đó khuyến nghị được sinh ra bởi HTVN sẽ có độ thỏa mãn cân bằng hơn giữa mọi người dùng trong nhóm.

Chương 3. HỆ TƯ VẤN NHÓM ĐỘNG THEO TIẾP CẬN TÍNH TOÁN MỜ ĐẢM BẢO TÍNH CÔNG BẰNG

3.1. Mở đầu

Trong luận án này, tập mờ trực cảm là trọng tâm nghiên cứu để xây dựng một biểu diễn phù hợp cho thông tin về đánh giá của người dùng với các sản phẩm, dịch vụ [55]–[57]. Bên cạnh đó, thông tin về đánh giá của người dùng được xem xét dưới tiếp cận thông tin động. Đây là hai yếu tố quan trọng để có thể xây dựng HTVN gần với thực tiễn hơn.

Trong chương ba này NCS trình bày nghiên cứu về “Hệ tư vấn nhóm động theo tiếp cận tính toán mờ đảm bảo tính công bằng”. Mục tiêu chính cần đạt được như sau:

- Nghiên cứu về Hệ tư vấn nhóm động để xây dựng phương pháp xử lý sự thay đổi về sở thích người dùng và sự suy hao về tính hấp dẫn của sản phẩm.

- Nghiên cứu Hệ tư vấn nhóm động dựa trên Tập mờ trực cảm để xử lý thông tin về sự do dự, không chắc chắn trong HTVN.

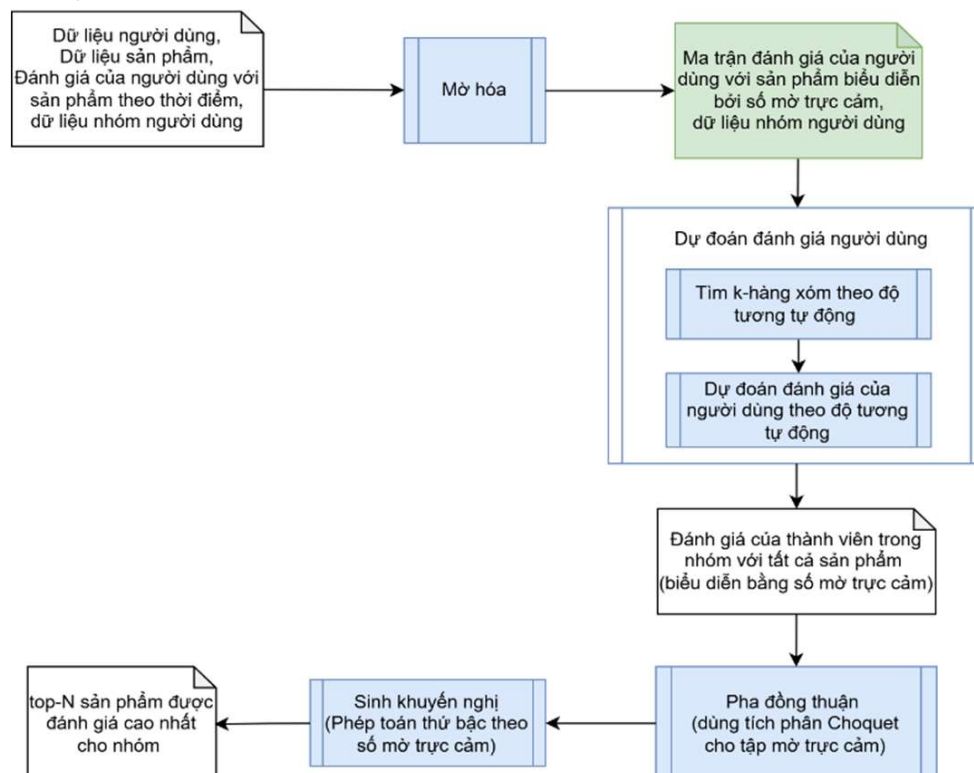
- Nghiên cứu và đề xuất mô hình HTVN động trên tập mờ trực cảm trong đó phép toán hợp sử dụng tích phân Choquet làm cơ sở để đảm bảo tính công bằng của kết quả khuyến nghị.

3.2. Đề xuất mô hình HTVN động trên tập mờ trực cảm

3.2.1. Đề xuất mô hình HTVN động trên tập mờ trực cảm:

a. Mô hình khái quát

Dựa trên tiếp cận xây dựng hệ tư vấn nhóm với pha đồng thuận được sử dụng để tạo ra khuyến nghị chung cho nhóm người dùng, mô hình hệ tư vấn nhóm động trên tập mờ trực cảm được thể hiện trong sơ đồ sau:



Hình 3.1 Sơ đồ mô hình Hệ tư vấn nhóm động trên tập mờ trực cảm

b. Các công thức được sử dụng trong mô hình

Bước 1: Mờ hóa các đánh giá

Bước 2: xác định độ tương tự động

Định nghĩa 3.1: cho t_{ui} , t_{uj} là thời điểm đánh giá của người dùng u_i , u_j đánh giá sản phẩm k . Ảnh hưởng của yếu tố thời gian lên độ tương tự của người dùng được đo bởi công thức:

$$f(\Delta t_i) = \frac{1}{1 + \lambda \Delta t} = \frac{1}{1 + \lambda |t_{ui} - t_{vi}|} \quad (3.4)$$

Độ tương tự động (Dynamic similarity):

$$Dsim(u, v) = Sim(u, v) \cdot f(\Delta t_i) \quad (3.5)$$

Bước 3: Dự đoán đánh giá của người dùng có xem xét tới ảnh hưởng của thời gian:

Để ước lượng đánh giá của một người dùng u cho sản phẩm j (biểu diễn trên số mờ trực cảm) ta có thể sử dụng một trong hai công thức sau.

$$a_{ui} = \bigoplus_{j=1}^n \frac{1}{n} a_{uj} \oplus IFAW(\tilde{A}_v, W_{DCosim(v, u)}) \quad (3.7)$$

$$a_{ui} = IFB^{p, q} a_u \oplus IFAW(\tilde{A}_v, W_{DCosim(v, u)}) \quad (3.8)$$

Trong đó:

$$IFAW(\tilde{A}_v, W_{DCosim(v, u)}) = \bigoplus_{v \in \tilde{U}} \tilde{a}_{vi} Dcosim(v, u) \quad (3.9)$$

3.2.2. Phép toán đồng thuận cho HTVN mờ trực cảm động

Tích phân Choquet cho tập mờ trực cảm:

Định nghĩa 3.3: Cho nhóm người dùng $g = \{u_1, u_2, \dots, u_n\}$, và véc tơ

$R_g = \left\{ \left(\mu_{u_i}, \nu_{u_i} \right); i = 1, 2, \dots, n \right\}$ là một tập các số mờ trực cảm thể hiện

đánh giá của các thành viên trong nhóm g cho một sản phẩm. Cho ξ là một hàm dung lượng trên tập người dùng g , ta có phép toán tổng hợp đánh giá của các thành viên dựa trên tích phân Choquet như sau:

$$IF_CA_{\xi}(R_g) = \left(1 - \prod_{i=1}^n (1 - \mu_{u_i})^{\xi(A_i^{\uparrow}) - \xi(A_{i+1}^{\uparrow})}, \prod_{i=1}^n \nu_{u_i}^{\xi(A_i^{\uparrow}) - \xi(A_{i+1}^{\uparrow})} \right) \quad (3.13)$$

Trong đó $A_i^{\uparrow}$ tập hợp các thành viên trong nhóm tương ứng, và

$A_i^{\uparrow} = \{u_i, u_{i+1}, \dots, u_n\}$ là các tập con của g , $A_{n+1}^{\uparrow} = \phi$.

3.2.3. Phương pháp học tham số của thuật toán:

Trong thuật toán DFGRS một số tham số sẽ ảnh hưởng tới hiệu quả việc dự đoán đánh giá của một người dùng với các sản phẩm, dịch vụ. Trong nghiên cứu này, thuật toán tối ưu hóa Bayes được sử dụng để học các tham số quan trọng đến hệ thống.

3.3. Thực nghiệm thuật toán

3.3.1. Dữ liệu thực nghiệm

Dữ liệu thực nghiệm của thuật toán hệ khuyến nghị nhóm mờ động được thực hiện trên cùng bộ dữ liệu MovieLens-1M đã được trình bày trong chương 2.

3.3.2. Phương thức đánh giá

Độ đo sự thỏa mãn trung bình của nhóm

Đánh giá của các thành viên đều thể hiện bởi một số mờ trực cảm nên giá trị thể hiện sự thỏa mãn trung bình của nhóm với tất cả các sản phẩm trong khuyến nghị I được tính bởi công thức sau.

$$g_pref(g, I) = \frac{1}{|I|} \sum_{i \in I} \left(\frac{\sum_{u \in g} utility_score(r_{ui})}{|g|} \right) \quad (3.27)$$

Các độ đo về tính công bằng

- Độ đo tỷ lệ hài lòng:

Thiết lập độ đo công bằng đầu tiên được sử dụng đó là tỷ lệ thành viên hài lòng với khuyến nghị trên tổng số người của nhóm.

$$fairness(g, i) = \frac{|\bigcup_{u \in g} r_{ui} \succ r_i^{th}|}{|g|} \quad (3.28)$$

Trong đó r_i^{th} là đánh giá trung bình của mỗi sản phẩm được tính bằng công thức sau.

$$r_i^{th} = IFAW(R_i, W) = \bigoplus_{u=1}^n r_{ui} w_u = \left\langle 1 - \prod_{u=1}^n (1 - \mu_{r_{ui}})^{w_u}, \prod_{u=1}^n \nu_{r_{ui}}^{w_u} \right\rangle \quad (3.29)$$

- **Độ đo công bằng:**

Độ đo thứ hai được sử dụng để đánh giá sự khác biệt trong mức độ thỏa mãn của từng thành viên nhóm. Độ đo này được gọi là độ đo công bằng, và được tính như sau.

$$equity(g, I) = \frac{1}{|I|} \sum_{i \in I} \left\{ \frac{1}{|g|^2} \sum_{l=1}^{|g|} \sum_{k=1}^{|g|} dis(r_{li}, r_{ki}) \right\} \quad (3.30)$$

- **Độ đo bất bình đẳng:**

Độ đo công bằng thứ ba là độ đo được đề xuất mới dựa trên khái niệm về chỉ số GINI thể hiện mức độ bất bình đẳng giữa một nhóm người dùng. Trong hệ tư vấn nhóm mờ động, độ đo bất bình đẳng được đề xuất và tính như sau:

$$Gini_equity(g, I) = \frac{\sum_{l=1}^{|g|} \sum_{k=1}^{|g|} |utility_score(r_{li}) - utility_score(r_{ki})|}{2|g| \sum_{u=1}^{|g|} utility_score(r_{ui})} \quad (3.31)$$

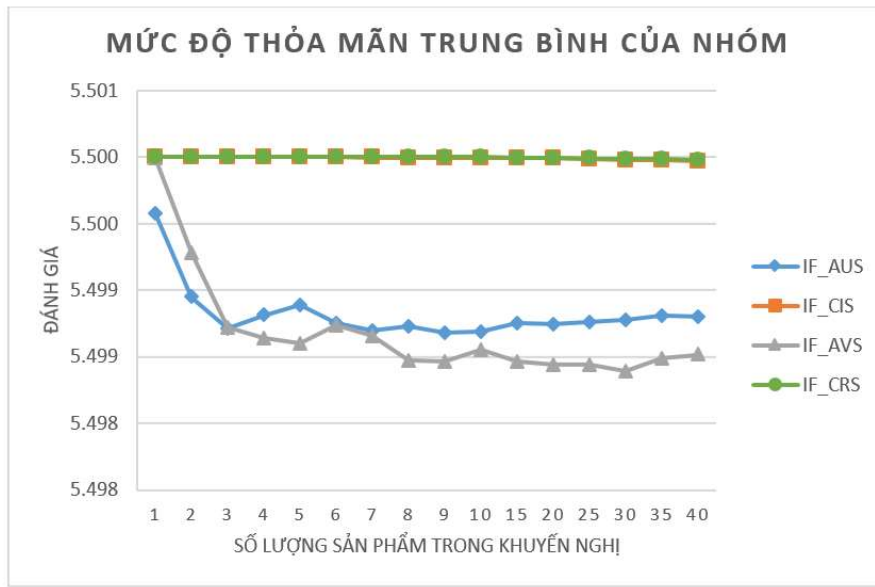
3.3.3. Kết quả và bàn luận

So sánh thuật toán

Trong phần này của luận án, hiệu quả của thuật toán đồng thuận sử dụng phép toán hợp dựa trên thích phân Choquet cho HTVN mờ động (với tên viết tắt là IF_CIS) sẽ được trình bày và so sánh với các chiến lược đồng thuận khác cho HTVN trên tập mờ trực cảm.

Kết quả so sánh bao gồm phân tích mức độ đánh giá trung bình của nhóm người dùng với khuyến nghị, mức độ công bằng của khuyến nghị theo các độ đo công bằng khác nhau. Dưới đây là tóm tắt một số kết quả so sánh chính.

Về mức độ thỏa mãn trung bình theo nhóm:



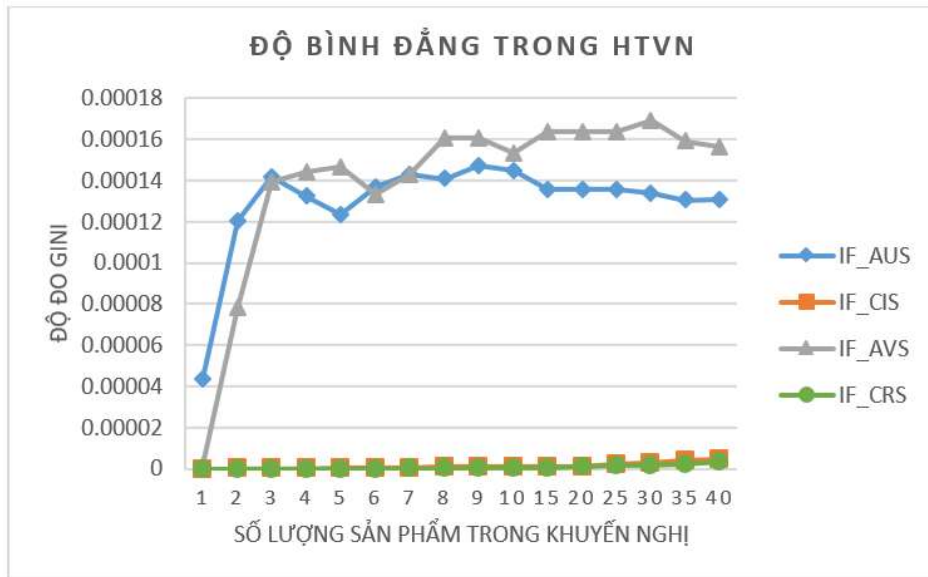
Hình 3.4 Độ thỏa mãn trung bình của nhóm theo bốn HTVN tốt nhất

Hình 3.4 cho thấy sự khác biệt nhỏ giữa nhóm bốn HTVN tốt hơn, trong đó thực nghiệm chứng tỏ rằng các HTVN mờ động sử dụng pha đồng thuận IF_CIS và IF_CRS có hiệu suất tốt hơn hai tiếp cận còn lại. Điều này cho thấy thuật toán sử dụng tích phân Choquet có hiệu quả tương đương với thuật toán tốt nhất về mặt độ thỏa mãn trung bình nhóm.

Về độ công bằng của khuyến nghị nhóm

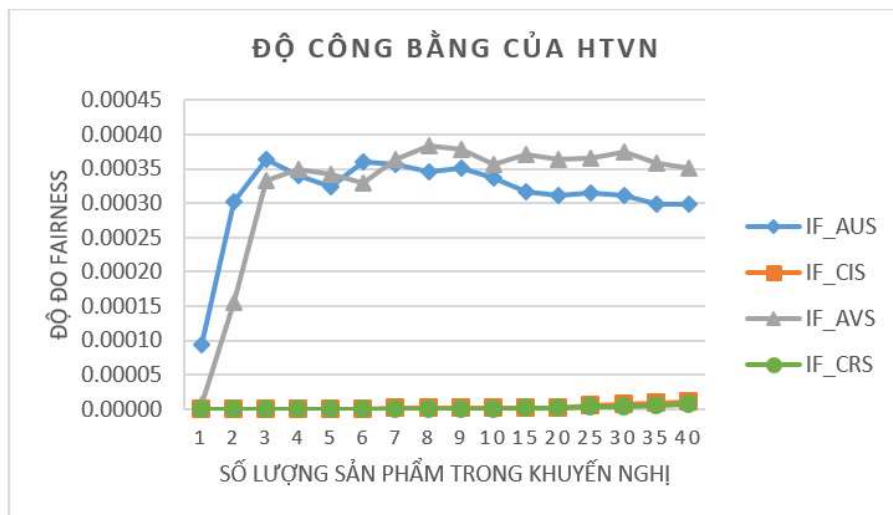
Hình 3.6 ở trên thể hiện hiệu quả của bốn thuật toán có ưu thế hơn về khía cạnh độ đo công bằng thứ hai. Trong bốn thuật toán này IF_AUS, IF_CIS, IF_AVS và IF_CRS, dễ thấy hai thuật toán IF_CIS và IF_CRS cho kết quả tốt hơn. Từ đó ta thấy thuật toán HTVN mờ

động sử dụng toán tử hợp với tích phân Choquet (IF_CIS) cho kết quả thuộc nhóm tốt nhất.



Hình 3.6 Độ công bằng của bốn HTVN mờ động tốt nhất

Xét theo độ công bằng GINI:



Hình 3.8 Độ công bằng GINI của bốn HTVN mờ động tốt nhất

Hình 3.8 ở trên thể hiện hiệu quả của bốn thuật toán có ưu thế hơn về khía cạnh độ đo công bằng thứ ba: độ bất bình đẳng. Trong bốn thuật toán này IF_AUS, IF_CIS, IF_AVS và IF_CRS, dễ thấy hai thuật toán IF_CIS và IF_CRS cho kết quả tốt hơn. Từ đó ta thấy thuật toán

HTVN mờ động sử dụng toán tử hợp với tích phân Choquet (IF_CIS) cho kết quả thuộc nhóm tốt nhất xét theo độ đo công bằng thứ 3.

3.4. Kết luận chương 3

Nghiên cứu xây dựng và phát triển hệ tư vấn nhóm động (HTVNĐ) dựa tập mờ trực cảm. Thông tin động được xử lý theo giả thiết sự ưa thích của người dùng, sự hấp dẫn của sản phẩm đều là thông tin biến đổi theo thời gian. Ngoài ra, trong pha động thuận của mô hình, toán tử Choquet vẫn được sử dụng để phát triển một phép toán hợp các đánh giá đơn của người dùng trong nhóm.

NHỮNG ĐÓNG GÓP MỚI CỦA LUẬN ÁN

Luận án nghiên cứu về Hệ tư vấn nhóm, một hệ tư vấn khái quát của hệ tư vấn đơn người dùng truyền thống, theo tiếp cận tính toán mờ trực cảm. Trong đó các thách thức của hệ tư vấn nhóm như tính công bằng của khuyến nghị, thông tin động, độ do dự và tính không chắc chắn của thông tin được nghiên cứu và đề xuất các hướng giải quyết. Dựa trên nghiên cứu lý thuyết và thực nghiệm, kết quả chính đạt được của luận án bao gồm:

1. Đề xuất thuật toán cho pha đồng thuận sử dụng tích phân Choquet áp dụng xây dựng hệ tư vấn nhóm nhằm tăng cường sự công bằng của các khuyến nghị và đảm bảo giữ được tổng thể lợi ích người dùng trong nhóm.

2. Đề xuất hệ tư vấn nhóm theo mô hình động sử dụng tích phân Choquet và tính toán mờ trực cảm trong pha đồng thuận đáp ứng giải quyết những bài toán có đánh giá của người dùng thay đổi theo thời gian.

Kết quả của luận án có thể được ứng dụng trong thực tế nhằm phát triển các hệ thống thông tin thông minh, cung cấp khả năng lọc thông tin phù hợp và đưa ra khuyến nghị cho nhóm người dùng.

DANH MỤC CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ

1. **Cu Nguyen Giap**, Nguyen Nhu Son*, Nguyen Long Giang, Hoang Thi Minh Chau, Tran Manh Tuan, Le Hoang Son, “A New Approach for Fairness Increment of Consensus-driven Group Recommender Systems Based on Choquet Integral”, International Journal of Data Warehousing and Mining (IJDWM), 2022, 18(1), 1-22. (ISSN 1548-3924, **SCIE**, **IF: 1,2**).
2. **Cu Nguyen Giap**, Le Thi Huyen Dieu, Luong Thi Hong Lan, Tran Thi Ngan, Tran Manh Tuan, “Utilize Deep learning to increase the performance of a Book recommender system using the Item-based Collaborative Filtering”. Hội thảo quốc tế RICE, 2022, 109-113.
3. Nguyễn Như Sơn, **Cù Nguyên Giáp**, Lê Hoàng Sơn , Nguyễn Long Giang, Dương Thị Thanh Loan, Trần Mạnh Tuấn, Dương Thị Thu Huyền, “Kỹ thuật tư vấn nhóm dựa trên tập mờ trực cảm và ứng dụng”, Hội thảo Quốc gia lần thứ XXV "Một số vấn đề chọn lọc của Công nghệ thông tin và Truyền thông" (2022), Hà Nội, Việt Nam, 2022, 81-86.
4. Nguyen Nhu Son, **Cu Nguyen Giap***, Le Hoang Son*, Nguyen Long Giang, Tran Manh Tuan, Vassilis C. Gerogiannis, Dimitrios Tzimos, “A Dynamic Fuzzy Group Recommender System based on Intuitionistic Fuzzy Choquet Integral Aggregation”, Soft Computing, 2024, 1-14. (E-ISSN 1433-7479, **SCIE**, **IF: 4,1**).