

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Trần Ngọc Khoa

**XÂY DỰNG HỆ HỖ TRỢ QUYẾT ĐỊNH TÌM CỐ VẤN HỌC TẬP CHO SINH VIÊN
DỰA TRÊN LÝ THUYẾT GHEP CẶP BỀN VỮNG VÀ CÁC GIẢI THUẬT TỐI ƯU
TIẾN HOÁ ĐA MỤC TIÊU**

Chuyên ngành: HỆ THỐNG THÔNG TIN

Mã số: 8.48.01.04 (Hệ thống thông tin)

TÓM TẮT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

HÀ NỘI - NĂM 2025

Đề án tốt nghiệp được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: PGS. TS. Lê Hải Châu

TS. Trịnh Bảo Ngọc

Phản biện 1:

Phản biện 2:

Đề án tốt nghiệp sẽ được bảo vệ trước Hội đồng chấm đề án tốt nghiệp thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: giờ ngày tháng năm

Có thể tìm hiểu đề án tốt nghiệp tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

MỞ ĐẦU

1. Lý do chọn đề tài

Trong bối cảnh giáo dục hiện đại, nhu cầu tiếp cận tri thức chuyên môn và sự hỗ trợ cá nhân hóa từ cổ vấn ngày càng trở nên cấp thiết. Điều này đặt ra yêu cầu cho việc nghiên cứu và phát triển các hệ thống ghép cặp sinh viên - cổ vấn hiệu quả và bền vững, nhằm tối ưu hóa quá trình học tập và nghiên cứu. Mặc dù việc ứng dụng công nghệ thông tin trong giáo dục là xu hướng tất yếu, thực trạng hiện nay cho thấy các hệ thống ghép cặp vẫn còn tồn tại nhiều hạn chế. Các hệ thống truyền thống thường thiếu tính cá nhân hóa, chưa đề cao việc xem xét sự phù hợp về chuyên môn, phương pháp giảng dạy và đặc điểm tính cách giữa sinh viên và cổ vấn. Điều này khiến sinh viên gặp khó khăn trong việc tìm kiếm người hướng dẫn phù hợp, ảnh hưởng tiêu cực đến hiệu quả học tập và nghiên cứu. Hơn nữa, việc ghép cặp thủ công cũng gây lãng phí thời gian và công sức của sinh viên, cổ vấn và các tổ chức giáo dục.

Để giải quyết những thách thức này, việc ứng dụng các thành tựu khoa học tiên tiến, điển hình là lý thuyết ghép cặp bền vững (Stable matching theory), vào thực tiễn giáo dục được xem là một hướng đi đầy tiềm năng. Được trao giải Nobel Kinh tế năm 2012, lý thuyết lĩnh vực lý thuyết trò chơi (Game theory) này đã tạo ra một bước đột phá trong việc giải quyết các bài toán phân bổ nguồn lực. Cụ thể, lý thuyết này cho phép tạo ra các cặp ghép đôi bền vững, trong đó không cá nhân nào có xu hướng rời bỏ cặp hiện tại để tìm kiếm đối tác khác phù hợp hơn.

Bằng cách kết hợp lý thuyết ghép cặp bền vững với các giải thuật tối ưu tiên hóa đa mục tiêu, đề án này hướng đến việc tối ưu hóa quy trình ghép cặp sinh viên - cổ vấn. Mục tiêu là xây dựng một hệ thống hỗ trợ quyết định thân thiện với người dùng, mang lại lợi ích tối đa cho cả sinh viên và cổ vấn. Việc ứng dụng thành tựu khoa học này vào thực tiễn hứa hẹn tạo ra một hệ thống ghép cặp hiệu quả, khoa học, với tiềm năng triển khai cao và khả năng mang lại lợi ích thiết thực cho cộng đồng giáo dục. Đề án được xây dựng trên nền tảng khoa học vững chắc, ứng dụng các giải thuật đã được nghiên cứu và kiểm chứng tính hiệu quả trong thực tế, từ đó gia tăng khả năng ứng dụng thực tiễn và đóng góp tích cực vào sự phát triển của lĩnh vực giáo dục.

2. Tổng quan về vấn đề nghiên cứu

Bài toán ghép cặp sinh viên - cố vấn đã và đang thu hút sự quan tâm nghiên cứu rộng rãi của cộng đồng khoa học, dẫn đến sự hình thành của nhiều phương pháp và kỹ thuật tiếp cận khác nhau. Các hệ thống dựa trên luật [14], tuy sở hữu ưu điểm về tính đơn giản và khả năng triển khai dễ dàng, lại thường thể hiện sự thiếu linh hoạt trong việc đáp ứng các yêu cầu phức tạp của bài toán thực tế. Trong khi đó, các nghiên cứu ứng dụng học máy [3] cho thấy tiềm năng đáng kể trong việc phân tích dữ liệu và tự động hóa quy trình ghép cặp. Tuy nhiên, các phương pháp này vẫn còn phụ thuộc vào chất lượng và số lượng dữ liệu huấn luyện, đồng thời gặp hạn chế trong việc diễn giải kết quả cho người dùng. Tương tự, các phương pháp tối ưu hóa dựa trên thống kê [7], mặc dù đạt hiệu quả trong việc tìm kiếm giải pháp tối ưu cho các hàm mục tiêu cụ thể, vẫn đối mặt với thách thức trong việc xử lý các bài toán đa ràng buộc phức tạp thường gặp trong thực tế. Các nghiên cứu đánh giá thực nghiệm [11], tuy cung cấp cái nhìn thực tiễn hơn về hiệu quả của các phương pháp cổ vấn hiện có, thường đòi hỏi chi phí đáng kể và khó khăn trong việc kiểm soát các yếu tố tác động.

Xuất phát từ những hạn chế của các phương pháp hiện tại, đề án này đề xuất ứng dụng Lý thuyết Ghép cặp Bền vững (Stable Matching Theory) – một lý thuyết đột phá được trao giải Nobel Kinh tế năm 2012 – làm nền tảng lý thuyết chính. Bên cạnh đó, đề án sẽ tập trung khai thác các giải thuật tối ưu hóa thuộc lớp Giải thuật Tiến hóa Đa mục tiêu (Multiobjective Evolutionary Algorithm - MOEA), đặc biệt chú trọng vào giải thuật NSGA-III, nhằm tìm kiếm lời giải tối ưu cho bài toán ghép cặp sinh viên - cố vấn học tập. Việc lựa chọn NSGA-III dựa trên khả năng xử lý hiệu quả các bài toán tối ưu đa mục tiêu với nhiều ràng buộc phức tạp. So sánh hiệu năng của NSGA-III với các giải thuật khác trong khuôn khổ MOEA, bao gồm eMOEA, VEGA, NSGA-II, IBEA, PESA2, sẽ cung cấp cái nhìn toàn diện về hiệu suất của từng giải thuật, hỗ trợ quá trình lựa chọn giải thuật phù hợp nhất cho bài toán.

Ý tưởng cốt lõi của đề án này là xây dựng một hệ thống hỗ trợ quyết định thông minh để khắc phục những yếu kém đó. Đề xuất trung tâm của đề án chính là giải thuật lai SMT-NSGA-III. Cách tiếp cận này khai thác sức mạnh của NSGA-III để khám phá không gian giải pháp rộng lớn, tìm ra một tập hợp các phương án ghép cặp ưu việt, cân bằng được nhiều mục tiêu (gọi là "mặt trận Pareto"), sau đó sử dụng các nguyên lý của SMT (thông qua giải thuật Gale-Shapley) như một bước tinh chỉnh cuối cùng để đảm bảo giải pháp được chọn là hoàn toàn bền vững. Sự kết hợp này giải quyết được những điểm yếu cố hữu khi sử dụng riêng lẻ từng lý thuyết: SMT gặp khó khăn với nhiều tiêu chí, trong khi MOEA không đảm bảo được tính bền vững.

Thông qua việc kết hợp ưu điểm của Lý thuyết Ghép cặp Bền vững với sức mạnh của các giải thuật MOEA, đề án hướng đến việc xây dựng một hệ thống hỗ trợ quyết định thông minh, linh hoạt và hiệu quả. Hệ thống này sẽ xem xét đồng thời nhiều yếu tố ảnh hưởng đến quyết định của cả sinh viên và cố vấn, bao gồm trình độ chuyên môn, kinh nghiệm, phong cách giảng dạy, mục tiêu học tập... Mục tiêu là đưa ra các giải pháp cân bằng mong muốn của cả hai bên, tối ưu hóa sự hài lòng và nâng cao hiệu quả của quá trình cố vấn. Với việc kết hợp những ưu điểm của các phương pháp hiện có và khắc phục những hạn chế còn tồn tại, đề án kỳ vọng sẽ mang đến một phương pháp tiếp cận toàn diện và hiệu quả hơn cho bài toán ghép cặp sinh viên - cố vấn học tập, đồng thời đóng góp vào việc nâng cao chất lượng giáo dục.

3. Mục đích nghiên cứu

Đề án hướng đến việc phát triển một hệ thống hỗ trợ quyết định thông minh và hiệu quả, tận dụng sức mạnh của lý thuyết ghép cặp bền vững (Stable matching theory) và các giải thuật tối ưu tiên hóa đa mục tiêu, nhằm giải quyết bài toán ghép cặp sinh viên - cố vấn học tập một cách tối ưu. Đề án đặt ra mục tiêu thiết kế và triển khai một hệ thống có khả năng phân tích, xử lý và đề xuất các cặp ghép đôi tối ưu, đáp ứng các tiêu chí và ràng buộc đa dạng từ cả phía sinh viên và cố vấn.

Cụ thể, các mục tiêu của đề án bao gồm:

- Phân tích bài toán ghép cặp: Nghiên cứu sâu rộng bài toán ghép cặp sinh viên - cố vấn học tập, xác định và phân tích các tiêu chí quan trọng ảnh hưởng đến quyết định lựa chọn cố vấn của sinh viên và ngược lại. Quá trình này bao gồm việc khảo sát, tổng hợp và đánh giá các yếu tố then chốt góp phần vào sự thành công của mối quan hệ cố vấn, từ đó xây dựng nền tảng cho việc phát triển mô hình và giải thuật.

- Phát triển mô hình và giải thuật: Đề xuất và phát triển một mô hình toán học mạnh mẽ, kết hợp với giải thuật tối ưu hóa quy trình ghép cặp, nhằm đảm bảo sự hài lòng và hiệu quả cho cả sinh viên và cố vấn. Mô hình này sẽ tích hợp các tiêu chí đã được xác định ở mục tiêu 1, đồng thời xem xét các ràng buộc thực tế để đảm bảo tính khả thi của giải pháp.

- Xây dựng hệ thống hỗ trợ quyết định: Xây dựng một hệ thống hỗ trợ quyết định dựa trên Lý thuyết Ghép cặp Bền vững, kết hợp giải thuật NSGA-III và giải thuật Gale-Shapley để hỗ trợ quá trình ghép cặp sinh viên - cố vấn. Việc tích hợp các giải thuật này nhằm tận dụng khả năng xử lý bài toán đa mục tiêu và đảm bảo tính bền vững của các cặp ghép đôi được đề xuất.

- Triển khai và đánh giá hệ thống: Triển khai và đánh giá hiệu quả của hệ thống thông qua mô phỏng và thực nghiệm. Quá trình đánh giá sẽ sử dụng các bộ dữ liệu mô phỏng và dữ liệu thực tế để kiểm chứng tính hiệu quả, khả năng mở rộng và tính ứng dụng thực tiễn của hệ thống. Kết quả đánh giá sẽ được phân tích và sử dụng để cải tiến và hoàn thiện hệ thống.

4. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu của đề án là bài toán ghép cặp sinh viên – cố vấn học tập trong lĩnh vực giáo dục - một bài toán phân bổ nguồn lực phức tạp đòi hỏi tính hiệu quả và sự hài lòng cao từ cả hai nhóm đối tượng: sinh viên với đa dạng nhu cầu, mục tiêu học tập và cố vấn học tập với khả năng, kinh nghiệm chuyên môn riêng biệt. Bài toán này đặt ra thách thức trong việc cân bằng đồng thời các yếu tố như chuyên môn phù hợp, thời gian cố vấn khả thi, khả năng đáp ứng về kinh phí, cũng như mức độ phù hợp về tính cách, phong cách làm việc giữa sinh viên và người cố vấn.

Phạm vi nghiên cứu của đề án bao gồm:

- Nghiên cứu các nguyên tắc của lý thuyết ghép cặp bền vững để xác định tiêu chí ghép sinh viên với cố vấn học tập và các yếu tố ảnh hưởng đến sự bền vững trong mối quan hệ giữa sinh viên và cố vấn.

- Nghiên cứu các giải thuật tối ưu hóa tiến hóa đa mục tiêu, thiết kế mô hình tối ưu hóa để đạt được sự phù hợp tối ưu giữa sinh viên và cố vấn dựa trên nhiều tiêu chí.

5. Phương pháp nghiên cứu

Phương pháp phân tích tổng kết kinh nghiệm

Thu thập, tổng hợp và phân tích các tài liệu và nghiên cứu trước đây về lý thuyết ghép cặp bền vững và giải thuật tối ưu đa mục tiêu, bài toán ghép cặp sinh viên - người cố vấn.

Phương pháp thực nghiệm

- Xây dựng hệ thống thực tế dựa trên mô hình và giải thuật đã thiết kế và chạy thử nghiệm hệ thống với dữ liệu mô phỏng.

- Sử dụng các tiêu chí định lượng (độ chính xác, thời gian chạy...) và định tính (khả năng sử dụng, sự hài lòng...) để đánh giá hiệu quả của hệ thống.

- So sánh kết quả với các phương pháp ghép nối truyền thống để chứng minh tính ưu việt của hệ thống.

CHƯƠNG 1 – MÔ TẢ BÀI TOÁN VÀ CƠ SỞ LÝ THUYẾT

1.1. Mô tả bài toán

1.1.1. Đặc điểm

Sự phức tạp của bài toán ghép cặp sinh viên và cố vấn học tập đến từ sự đa dạng và không đồng nhất của các bên tham gia. Về phía sinh viên, mỗi người là một cá nhân độc đáo với các nhu cầu học thuật, mục tiêu nghề nghiệp, định hướng nghiên cứu và phong cách học tập riêng biệt. Họ không chỉ cần người có chuyên môn mà còn cần một người thầy tận tâm, truyền cảm hứng. Tương ứng, các cố vấn học tập cũng có lĩnh vực chuyên môn, kinh nghiệm và phương pháp sư phạm khác nhau, đồng thời bị giới hạn về số lượng sinh viên họ có thể hướng dẫn một cách hiệu quả (sức chứa). Chính vì sự phức tạp này, các phương pháp ghép cặp truyền thống ngày càng bộc lộ nhiều hạn chế. Việc ghép cặp thường được thực hiện thủ công, dựa trên kinh nghiệm chủ quan hoặc các tiêu chí đơn giản, dẫn đến thiếu tính cá nhân hóa và không chú trọng đến sự phù hợp sâu sắc về chuyên môn hay tính cách. Hệ thống này không chỉ tốn thời gian và công sức mà còn dễ tạo ra các quyết định không tối ưu, thiếu công bằng, làm ảnh hưởng đến hiệu quả học tập của sinh viên và bỏ lỡ những kết nối tiềm năng mang lại lợi ích lớn cho cả hai phía..

1.1.2. Yêu cầu

Để giải quyết bài toán ghép cặp sinh viên - cố vấn một cách hiệu quả, một hệ thống hỗ trợ quyết định tiên tiến cần phải đáp ứng nhiều yêu cầu cốt lõi. Trước hết, hệ thống phải có khả năng xem xét đa tiêu chí một cách toàn diện, cân nhắc đồng thời các yếu tố phức tạp từ cả hai phía như nguyện vọng học thuật của sinh viên và chuyên môn, sức chứa của cố vấn. Mục tiêu không chỉ là tìm ra một cặp ghép phù hợp, mà phải tối ưu hóa sự hài lòng, cân bằng lợi ích đa chiều và đảm bảo các mối quan hệ này bền vững, ổn định lâu dài, tránh tình trạng các bên muốn phá vỡ để tìm đối tác tốt hơn.

Bên cạnh đó, quy trình ghép cặp phải được xây dựng trên nền tảng khoa học, khách quan và minh bạch để tạo dựng niềm tin, lý tưởng nhất là có thể giải thích được các quyết định đề xuất. Cuối cùng, một yêu cầu thực tiễn không thể thiếu là hệ thống phải có giao diện thân thiện, trực quan và dễ sử dụng cho mọi đối tượng người dùng (sinh viên, cố vấn, nhà quản lý), đảm bảo công nghệ là công cụ hỗ trợ đắc lực thay vì trở thành một rào cản phức tạp..

1.2. Lý thuyết ghép cặp bền vững (Stable matching theory)

1.2.1. Khái niệm

Lý thuyết ghép cặp bền vững là một nhánh quan trọng của lý thuyết trò chơi, tập trung vào việc tạo ra các cặp ghép "ổn định" giữa hai nhóm đối tượng. Cốt lõi của lý thuyết này là khái niệm "ổn định", mô tả một trạng thái mà không có bất kỳ hai người nào từ hai nhóm khác nhau có động lực rời bỏ đối tác hiện tại để hình thành một cặp mới với nhau. Về cơ bản, một cặp ghép ổn định là cặp ghép không chứa bất kỳ "cặp chặn" (blocking pair) nào. Lý thuyết này được đặt nền móng bởi công trình của Gale và Shapley vào năm 1962 với giải thuật nổi tiếng của họ.

Bài toán hôn nhân ổn định là ví dụ kinh điển để minh họa. Trong bài toán này, mỗi người trong hai nhóm (ví dụ: nam và nữ) có một danh sách ưu tiên về các đối tác tiềm năng. Một cặp ghép sẽ bị coi là không ổn định nếu tồn tại một "cặp chặn" - tức là một người nam và một người nữ không được ghép với nhau nhưng cả hai đều thích đối phương hơn người bạn đời hiện tại của mình. Sự tồn tại của cặp chặn này tạo ra động lực để họ phá vỡ mối quan hệ được gán, gây ra sự bất ổn cho toàn bộ hệ thống ghép cặp.).

1.2.2. Các mô hình ghép cặp

Lý thuyết ghép cặp bền vững đã phát triển từ mô hình một-một nền tảng, với ví dụ kinh điển là bài toán hôn nhân ổn định, sang các dạng phức tạp hơn để giải quyết nhiều vấn đề thực tế. Một trong những mở rộng quan trọng là mô hình ghép cặp nhiều-một, thường được áp dụng trong tuyển sinh đại học. Trong mô hình này, một bên (như trường đại học) có thể ghép với nhiều đối tượng (sinh viên) theo một chỉ tiêu (quota) nhất định, trong khi mỗi sinh viên chỉ có thể theo học tại một trường duy nhất.

Mô hình tổng quát nhất là ghép cặp nhiều-nhiều, cho phép mỗi bên đều có thể kết nối với nhiều đối tác, ví dụ như nhân viên tham gia nhiều dự án và mỗi dự án có nhiều nhân viên. Ngoài ra, lý thuyết này còn được mở rộng với các biến thể như ghép cặp có trao đổi giá trị (chuyển nhượng) hoặc có các ràng buộc bổ sung. Sự đa dạng của các mô hình này cho thấy tính linh hoạt và khả năng ứng dụng mạnh mẽ của lý thuyết ghép cặp bền vững trong việc giải quyết các bài toán phân bổ phức tạp trong thế giới thực..

1.2.3. Sự ổn định trong ghép cặp

Khái niệm "ổn định" (stability) là yếu tố then chốt và là thước đo chất lượng trong lý thuyết ghép cặp bền vững. Một cặp ghép được coi là ổn định khi không tồn tại bất kỳ "cặp chặn" (blocking pair) nào – tức là không có hai đối tượng nào từ hai nhóm khác nhau muốn rời bỏ đối tác hiện tại của mình để hình thành một cặp mới tốt hơn cho cả hai. Điều này nhằm đảm bảo sự hài lòng của các bên và ngăn chặn sự xáo trộn sau khi việc ghép cặp hoàn tất. Tuy nhiên, định nghĩa chính xác của sự ổn định thay đổi và trở nên phức tạp hơn tùy thuộc vào từng mô hình ghép cặp cụ thể. Trong mô hình một-một cơ bản, sự ổn định chỉ đơn giản là không có cặp nào thích nhau hơn đối tác hiện tại của họ. Với mô hình nhiều-một (ví dụ: tuyển sinh đại học), sự bất ổn xảy ra khi một sinh viên thích một trường hơn, và trường đó hoặc còn chỉ tiêu tuyển sinh, hoặc sẵn sàng loại một sinh viên hiện tại để nhận sinh viên mới này. Ở mức độ phức tạp nhất là mô hình nhiều-nhiều, các khái niệm ổn định còn được mở rộng hơn nữa, chẳng hạn như "ổn định nhóm" – một điều kiện chặt chẽ hơn, yêu cầu không có bất kỳ một nhóm đối tượng nào có thể cùng nhau tái sắp xếp để tất cả đều có lợi hơn.

1.2.4. Giải thuật Gale-Shapley

Giải thuật Gale-Shapley là một phương pháp kinh điển, có tầm quan trọng được công nhận qua giải Nobel Kinh tế, dùng để giải quyết bài toán ghép cặp bền vững (dạng một-một và nhiều-một). Giải thuật hoạt động dựa trên một quy trình lặp đi lặp lại: một bên (bên "cầu hôn") sẽ lần lượt đưa ra lời đề nghị cho các đối tác theo thứ tự ưu tiên của mình. Bên còn lại (bên "nhận lời") sẽ tạm thời chấp nhận lời đề nghị tốt nhất mà họ nhận được, và có thể từ chối đối tác hiện tại nếu một lời đề nghị hấp dẫn hơn xuất hiện. Quá trình này tiếp tục cho đến khi một trạng thái ổn định, không còn sự thay đổi nào, được thiết lập.

Giải thuật này có những đặc tính toán học quan trọng: nó luôn dừng lại, luôn đảm bảo tạo ra một cặp ghép bền vững (không có cặp chặn), và đặc biệt là tối ưu cho bên đưa ra lời đề nghị. Điều này có nghĩa là bên cầu hôn sẽ nhận được đối tác tốt nhất có thể trong số tất cả các kết quả bền vững, nhưng ngược lại, kết quả này có thể là tệ nhất cho bên nhận lời. Tuy nhiên, giải thuật cũng có hạn chế khi kết quả của nó phụ thuộc nhiều vào việc bên nào được chọn làm bên đề xuất và không thể cân bằng đồng thời nhiều tiêu chí phức tạp khác ngoài danh sách ưu tiên, chẳng hạn như sự cân bằng khối lượng công việc hay sự tương thích về tính cách.

1.3. Giải thuật tối ưu tiến hoá đa mục tiêu (MOEA)

1.3.1. Khái niệm

Các Giải thuật Tiến hóa Đa mục tiêu (MOEA) là các giải thuật tìm kiếm lấy cảm hứng từ quá trình tiến hóa tự nhiên (chọn lọc, lai ghép, đột biến) để giải quyết các bài toán có nhiều mục tiêu thường xung đột với nhau. Khác biệt cơ bản so với các phương pháp truyền thống chỉ tìm một giải pháp duy nhất, mục tiêu của MOEA là xác định một tập hợp các giải pháp tối ưu, gọi là "tập Pareto tối ưu". Mỗi giải pháp trong tập này đại diện cho một sự đánh đổi khác nhau, nơi không có giải pháp nào tốt hơn hẳn các giải pháp khác ở tất cả mọi mục tiêu.

Khái niệm trung tâm của phương pháp này là sự "trội Pareto": một giải pháp được xem là trội hơn giải pháp khác nếu nó tốt hơn ở ít nhất một mục tiêu mà không tệ hơn ở bất kỳ mục tiêu nào còn lại. Một giải pháp được gọi là Pareto tối ưu khi không có giải pháp nào khác có thể trội hơn nó. Mục tiêu cuối cùng của một giải thuật MOEA là tìm ra một tập hợp giải pháp, được gọi là "mặt trận Pareto", để xấp xỉ tốt nhất cho tập hợp các lựa chọn tối ưu, đảm bảo các giải pháp vừa hiệu quả vừa đa dạng. Điều này cung cấp cho người ra quyết định một loạt các phương án cân bằng khác nhau để lựa chọn, thay vì chỉ một câu trả lời duy nhất.

1.3.2. Các giải thuật MOEA phổ biến

Ba giải thuật tiến hóa đa mục tiêu (MOEA) phổ biến được giới thiệu: NSGA-II, SPEA2, và MOEA/D. NSGA-II nổi bật với hai cơ chế chính: "phân loại không thống trị nhanh" để xếp hạng các giải pháp thành các mặt trận Pareto khác nhau và "khoảng cách dồn ép" để duy trì sự đa dạng bằng cách ưu tiên các giải pháp ở những vùng thưa thớt hơn trên mặt trận. SPEA2 thì sử dụng một kho lưu trữ ngoài để lưu giữ các giải pháp tốt nhất đã tìm được, kết hợp việc đánh giá "sức mạnh" (số lượng giải pháp mà một giải pháp thống trị) và "mật độ" (đo lường sự đông đúc xung quanh) để vừa tìm kiếm hiệu quả vừa đảm bảo sự đa dạng.

Khác biệt với hai giải thuật trên, MOEA/D áp dụng chiến lược phân rã, biến một bài toán đa mục tiêu phức tạp thành một tập hợp các bài toán con đơn mục tiêu, mỗi bài toán tương ứng với một vector trọng số (hướng ưu tiên). Giải thuật tối ưu hóa đồng thời các bài toán con này và tận dụng thông tin từ các "bài toán hàng xóm" để tăng tốc độ hội tụ. Nhìn chung, cả ba giải thuật đều nhằm mục đích tìm ra một tập hợp các giải pháp tối ưu (mặt trận Pareto) đa dạng và hiệu quả, nhưng mỗi loại sử dụng một chiến lược riêng để cân bằng giữa việc tìm kiếm giải pháp tốt và duy trì sự đa dạng của chúng.

1.3.3. Giải thuật NSGA-III

NSGA-III là một phiên bản cải tiến của giải thuật NSGA-II, được thiết kế đặc biệt để giải quyết các bài toán tối ưu hóa có nhiều mục tiêu (từ bốn mục tiêu trở lên). Trong những bài toán này, các phương pháp truyền thống như của NSGA-II gặp khó khăn vì phần lớn các giải pháp đều trở nên "không bị trội" lẫn nhau, làm giảm hiệu quả của việc chọn lọc. Hơn nữa, cơ chế đo "khoảng cách dồn ép" để duy trì sự đa dạng cũng trở nên kém hiệu quả trong không gian đa chiều. Để khắc phục điều này, NSGA-III giới thiệu một cơ chế chọn lọc đột phá dựa trên các "điểm tham chiếu" được xác định trước, giúp định hướng quá trình tìm kiếm một cách có hệ thống.

Trong quá trình hoạt động, sau khi sắp xếp các giải pháp thành các mặt trận không bị trội như NSGA-II, NSGA-III sử dụng các điểm tham chiếu để thực hiện bước chọn lọc quyết định. Mỗi giải pháp được liên kết với một điểm tham chiếu gần nhất. Để duy trì sự đa dạng, giải thuật sẽ ưu tiên giữ lại các giải pháp thuộc về những điểm tham chiếu đang có ít giải pháp liên kết nhất. Nếu một điểm tham chiếu có nhiều ứng cử viên, giải pháp ở gần đường tham chiếu nhất sẽ được chọn để thúc đẩy sự hội tụ. Bằng cách này, NSGA-III có khả năng duy trì sự đa dạng của các giải pháp một cách hiệu quả trong không gian mục tiêu có số chiều cao, đồng thời đảm bảo sự hội tụ mạnh mẽ về phía mặt trận Pareto thực sự.

1.4. Các nghiên cứu liên quan

Các phương pháp phân công cố vấn học tập truyền thống thường là thủ công, dựa trên kinh nghiệm hoặc các tiêu chí đơn giản, dẫn đến sự không phù hợp và thiếu hài lòng. Các nỗ lực cải tiến hiện tại như các hệ thống quản lý gia sư hay ứng dụng học máy, dù mang lại sự tiện lợi và cá nhân hóa, vẫn còn nhiều hạn chế. Các hệ thống này thường không giải quyết được bài toán ghép cặp hai phía một cách toàn diện, phụ thuộc nhiều vào dữ liệu, và không đảm bảo được các ràng buộc quan trọng như sức chứa của cố vấn hay tính "bền vững" của cặp ghép theo lý thuyết. Ngay cả giải thuật kinh điển Gale-Shapley cũng không đủ mạnh vì chỉ xét được một tiêu chí ưu tiên duy nhất.

Để khắc phục những hạn chế này, hướng đi hiệu quả là sử dụng các Giải thuật Tiến hóa Đa mục tiêu (MOEA) như NSGA-II, SPEA2, hay MOEA/D, bởi chúng có khả năng tối ưu hóa đồng thời nhiều mục tiêu phức tạp và xung đột của cả sinh viên và cố vấn. Mặc dù vậy, các nghiên cứu hiện tại vẫn còn tồn tại một số điểm yếu như nguồn dữ liệu hạn hẹp và bộ tiêu chí đánh giá chưa đầy đủ. Do đó, giải pháp tối ưu là xây dựng một hệ thống hỗ trợ

quyết định kết hợp nền tảng của lý thuyết ghép cặp bền vững với sức mạnh của các giải thuật MOEA, đồng thời phải dựa trên một bộ tiêu chí đánh giá sự phù hợp toàn diện và chính xác hơn.

CHƯƠNG 2 – XÂY DỰNG GIẢI THUẬT VÀ HỆ TRỢ GIÚP QUYẾT ĐỊNH

2.1. Xây dựng mô hình toán học

2.1.1. Tập hợp các đối tượng

Bài toán ghép cặp này xác định hai nhóm đối tượng chính: sinh viên và cố vấn học tập. Mỗi sinh viên được mô tả bởi một tập hợp thuộc tính đa dạng, không chỉ giới hạn ở năng lực học thuật mà còn bao gồm các yếu tố như điểm trung bình (GPA), lĩnh vực quan tâm, kỹ năng, mục tiêu nghề nghiệp và phong cách học tập. Một điểm quan trọng là thay vì sử dụng một danh sách ưu tiên tường minh, sự phù hợp và sở thích của sinh viên được thể hiện một cách ngầm định thông qua chính các thuộc tính này.

Tương tự, mỗi cố vấn học tập được định nghĩa bởi các thuộc tính chuyên môn như lĩnh vực chuyên sâu, kinh nghiệm, phương pháp giảng dạy, và đặc biệt là một ràng buộc cứng về "sức chứa" (capacity) - số lượng sinh viên tối đa họ có thể hướng dẫn. Sở thích của họ đối với sinh viên cũng được xác định qua một bộ tiêu chí thay vì một danh sách xếp hạng đơn thuần. Chất lượng và tính đầy đủ của dữ liệu đầu vào này là yếu tố nền tảng, ảnh hưởng trực tiếp đến độ tin cậy và hiệu quả của các quyết định ghép cặp mà hệ thống đưa ra.

2.1.2. Xác định các tiêu chí

Việc xác định các tiêu chí ghép cặp đóng vai trò then chốt, không phải để người dùng tạo ra danh sách ưu tiên thủ công, mà là để hệ thống tự động đánh giá mức độ phù hợp. Các tiêu chí này được xem xét từ ba góc độ khác nhau. Từ phía sinh viên, các yếu tố quan trọng bao gồm sự phù hợp về chuyên môn, sự tương thích giữa phong cách học tập và phương pháp giảng dạy, cũng như kinh nghiệm và sự tận tâm của cố vấn. Ngược lại, từ phía cố vấn, họ thường quan tâm đến tiềm năng học thuật, sự chủ động và sự tương đồng trong định hướng nghiên cứu của sinh viên.

Bên cạnh đó, mô hình còn phải đáp ứng các mục tiêu của cơ sở giáo dục, như đảm bảo tất cả sinh viên đều được hỗ trợ và quan trọng là phải cân bằng khối lượng công việc giữa các

cổ vấn. Tất cả các tiêu chí này sẽ được sử dụng làm nền tảng để hệ thống tính toán các "điểm tương thích" giữa mỗi cặp sinh viên-cổ vấn, từ đó xây dựng các hàm mục tiêu cho giải thuật tối ưu hóa, hướng đến việc tìm ra giải pháp ghép cặp khách quan và hiệu quả nhất cho toàn hệ thống.

2.1.3. Xây dựng các hàm mục tiêu

Để giải quyết bài toán ghép cặp, mô hình được xây dựng như một bài toán tối ưu đa mục tiêu, trong đó "sự ưu tiên" được thể hiện ngầm định thông qua các điểm tương thích được tính toán. Các hàm mục tiêu chính được thiết kế để cân bằng nhiều yếu tố cùng lúc. Các mục tiêu này bao gồm việc tối đa hóa sự hài lòng chung cho cả sinh viên (f_1) và cổ vấn học tập (f_2) bằng cách tổng hợp các điểm tương thích từ góc nhìn của mỗi bên.

Bên cạnh sự hài lòng cá nhân, mô hình còn hướng đến các mục tiêu mang tính hệ thống. Cụ thể là tối thiểu hóa sự mất cân bằng trong khối lượng công việc giữa các cổ vấn (f_3) để đảm bảo sự công bằng và chất lượng. Quan trọng nhất là mục tiêu tối thiểu hóa số lượng "cặp chẵn" (f_5), một khái niệm từ lý thuyết ghép cặp bền vững, để đảm bảo giải pháp cuối cùng ổn định và không ai có động lực phá vỡ cặp ghép của mình. Tất cả các mục tiêu này sẽ được tích hợp vào giải thuật NSGA-III để tìm ra một tập hợp các giải pháp tối ưu, mỗi giải pháp thể hiện một sự đánh đổi khác nhau giữa các tiêu chí.

2.1.4. Các ràng buộc trong bài toán

Để một giải pháp ghép cặp sinh viên - cổ vấn được coi là hợp lệ và có thể triển khai trong thực tế, nó phải tuân thủ một tập hợp các ràng buộc cơ bản. Các ràng buộc này định hình không gian các giải pháp khả thi, đảm bảo kết quả cuối cùng có ý nghĩa và phù hợp với môi trường giáo dục.

Hai ràng buộc quan trọng nhất được xác định là: thứ nhất, mỗi sinh viên chỉ được phép ghép cặp với tối đa một cổ vấn học tập tại một thời điểm, đảm bảo mối quan hệ một-một từ phía sinh viên. Thứ hai, và cũng là một ràng buộc cứng, mỗi cổ vấn học tập có một "sức chứa" (capacity) nhất định, tức là tổng số sinh viên được phân công cho một cổ vấn không được vượt quá giới hạn này. Việc tuân thủ nghiêm ngặt các ràng buộc này là điều bắt buộc, giúp ngăn ngừa tình trạng quá tải cho cổ vấn và đảm bảo tính hợp lý của toàn bộ hệ thống phân công..

2.2. Thiết kế giải thuật đề xuất

2.2.1. Ý tưởng kết hợp giải thuật NSGA-III và Gale-Shapley

Ý tưởng cốt lõi của giải thuật lai ghép là giải quyết bài toán ghép cặp khi không có danh sách ưu tiên tường minh từ người dùng. Thay vào đó, hệ thống sử dụng các "điểm tương thích" được tính toán một cách khách quan dựa trên thuộc tính đa chiều của sinh viên và cố vấn.

NSGA-III đóng vai trò là khung tối ưu hóa chính, có nhiệm vụ khám phá không gian giải pháp để tìm ra một tập hợp các phương án ghép cặp ưu việt (mặt trận Pareto). Giải thuật này cân bằng đồng thời nhiều mục tiêu: tối đa hóa sự hài lòng của sinh viên và cố vấn, cân bằng khối lượng công việc, và tối thiểu hóa số lượng "cặp chặn" - tất cả đều được tính toán dựa trên điểm tương thích.

Sau khi NSGA-III hoàn thành, một giai đoạn tinh chỉnh cuối cùng sử dụng giải thuật Gale-Shapley được thực hiện. Ở bước này, các danh sách ưu tiên cần thiết cho Gale-Shapley được tạo ra một cách tự động bằng cách sắp xếp tất cả các đối tác tiềm năng dựa trên điểm tương thích của họ. Mục đích của bước này là để đảm bảo rằng giải pháp cuối cùng thu được sẽ hoàn toàn bền vững theo định nghĩa của lý thuyết ghép cặp bền vững, loại bỏ triệt để bất kỳ sự bất ổn nào còn sót lại.

Cách tiếp cận này tận dụng sức mạnh của cả hai phương pháp: khả năng xử lý các đánh đổi phức tạp của NSGA-III và sự đảm bảo chặt chẽ về tính ổn định của Gale-Shapley, nhằm tạo ra các giải pháp vừa tối ưu đa mục tiêu, vừa ổn định..

2.2.2. Các bước thực hiện

Giải thuật lai ghép SMT-NSGA-III được đề xuất hoạt động theo một quy trình có cấu trúc, bắt đầu bằng việc chuẩn bị dữ liệu đầu vào, bao gồm thuộc tính của sinh viên, cố vấn, và quan trọng nhất là các hàm để tính toán "điểm tương thích" hai chiều, thay thế cho danh sách ưu tiên truyền thống. Mỗi giải pháp ghép cặp tiềm năng được mã hóa thành một "cá thể" trong quần thể của giải thuật. Giai đoạn chính của quy trình là vòng lặp tiến hóa của NSGA-III, trong đó mỗi giải pháp được đánh giá dựa trên một tập hợp các hàm mục tiêu đa dạng: tối đa hóa sự hài lòng của sinh viên và cố vấn (dựa trên điểm tương thích), cân bằng khối lượng công việc, và tối thiểu hóa số lượng "cặp chặn" được định nghĩa lại theo điểm tương thích. Sau khi NSGA-III đã chạy và tạo ra một tập hợp các giải pháp cân bằng, chất lượng cao (mặt

trận Pareto), giải thuật sẽ bước vào giai đoạn "tinh chỉnh" cuối cùng. Ở bước này, từ các điểm tương thích đã tính, hệ thống sẽ tự động tạo ra các danh sách ưu tiên có thứ tự cho cả sinh viên và cố vấn. Các danh sách này sau đó được sử dụng làm đầu vào cho giải thuật Gale-Shapley kinh điển để đảm bảo rằng giải pháp cuối cùng là hoàn toàn "bền vững", loại bỏ triệt để bất kỳ cặp chặn nào còn sót lại. Quá trình hai giai đoạn này kết hợp sức mạnh khám phá đa mục tiêu của NSGA-III với sự đảm bảo về tính ổn định của lý thuyết ghép cặp bền vững.

2.2.3. Phân tích độ phức tạp

Độ phức tạp tính toán của giải thuật lai ghép SMT-NSGA-III được xác định bởi ba giai đoạn chính. Đầu tiên là chi phí tính toán ban đầu để xác định tất cả các "điểm tương thích" giữa mỗi sinh viên và cố vấn, một bước có thể tốn kém nhưng thường chỉ thực hiện một lần. Giai đoạn thứ hai, và cũng là phần nặng nề nhất, là vòng lặp tiến hóa của NSGA-III. Trong mỗi thế hệ, mặc dù hầu hết các hàm mục tiêu (như sự hài lòng hay cân bằng tải) được tính toán tương đối nhanh, thì việc tính toán hàm mục tiêu f_5 (số lượng "cặp chặn" để đo lường tính bền vững) lại đặc biệt phức tạp, vì nó đòi hỏi phải duyệt qua một số lượng lớn các cặp chưa được ghép. Giai đoạn cuối cùng là bước tinh chỉnh tùy chọn bằng giải thuật Gale-Shapley. Giai đoạn này có chi phí riêng, bao gồm việc tạo ra các danh sách ưu tiên một cách tự động (thông qua việc sắp xếp dựa trên điểm tương thích) và sau đó là thực thi chính giải thuật Gale-Shapley. Tóm lại, việc không dùng danh sách ưu tiên có sẵn đã chuyển gánh nặng tính toán sang việc xác định điểm tương thích và, quan trọng hơn, là việc liên tục kiểm tra tính ổn định (số cặp chặn) trong suốt quá trình tối ưu hóa, đây chính là yếu tố chi phối độ phức tạp tổng thể của giải thuật.

2.2.4. Minh họa giải thuật

Phần này minh họa cơ chế hoạt động của giải thuật bằng một ví dụ cụ thể, thay thế danh sách ưu tiên tường minh bằng các "điểm tương thích" được tính toán. Dựa trên một giải pháp ghép cặp mẫu, ví dụ cho thấy cách hệ thống đánh giá chất lượng bằng việc tính toán các hàm mục tiêu. Cụ thể, sự hài lòng của sinh viên (f_1) và cố vấn (f_2) được định lượng bằng cách cộng tổng các điểm tương thích tương ứng, trong khi sự cân bằng tải (f_3) được đo lường bằng phương sai, tất cả đều dựa trên dữ liệu khách quan.

Ví dụ cũng làm rõ cách kiểm tra tính bền vững (f_5), trong đó một "cặp chặn" chỉ tồn tại khi cả sinh viên và cố vấn đều có điểm tương thích với nhau cao hơn so với đối tác hiện

tại của họ. Cuối cùng, phần này mô tả giai đoạn tinh chỉnh tùy chọn: sau khi NSGA-III tìm ra các giải pháp tốt, các danh sách ưu tiên có thể được tạo ra một cách tự động bằng cách sắp xếp tất cả các đối tác tiềm năng dựa trên điểm tương thích. Việc áp dụng giải thuật Gale-Shapley lên các danh sách động này hoạt động như một bước "đánh bóng", đảm bảo giải pháp cuối cùng là hoàn toàn bền vững, kết hợp được lợi thế của cả hai phương pháp.

2.3. Triển khai hệ trợ giúp quyết định

2.3.1. Kiến trúc hệ thống

Hệ thống được xây dựng theo kiến trúc client-server, phân tách rõ ràng giữa giao diện người dùng (frontend) và phần xử lý logic (backend), giúp dễ dàng bảo trì và mở rộng. Phần frontend, được phát triển bằng ReactJS, chịu trách nhiệm cung cấp một giao diện người dùng hiện đại và mượt mà, cho phép người dùng tải lên dữ liệu (file Excel), gửi yêu cầu và nhận, trực quan hóa kết quả phân tích.

Trong khi đó, phần backend là trung tâm xử lý của hệ thống, được xây dựng trên nền tảng Java 17, Spring Boot và MOEA Framework. Sự kết hợp này đảm bảo hiệu năng và sự ổn định, đồng thời cung cấp sẵn các giải thuật tối ưu hóa đa mục tiêu cần thiết cho bài toán ghép cặp. Backend có nhiệm vụ nhận dữ liệu, xử lý các file đầu vào, thực thi giải thuật ghép cặp phức tạp, và trả về kết quả cho frontend hiển thị.

2.3.2. Công nghệ sử dụng

Việc lựa chọn công nghệ được dựa trên các tiêu chí về hiệu năng, khả năng mở rộng, tính ổn định, bảo mật, cộng đồng hỗ trợ, và tính phù hợp với bài toán:

Backend: Java 17 được chọn làm ngôn ngữ lập trình chính do tính mạnh mẽ, nền tảng và là phiên bản LTS. Spring Boot được sử dụng làm framework chính, giúp đơn giản hóa việc phát triển ứng dụng Java, cung cấp tính module hóa cao và tích hợp liền mạch với các công nghệ khác. MOEA Framework, một thư viện Java chuyên dụng, được sử dụng để thực hiện các giải thuật tối ưu tiên hóa đa mục tiêu.

Frontend: React là thư viện JavaScript phổ biến, cho phép xây dựng UI hiện đại, tương tác cao và dễ bảo trì. React cho phép xây dựng các component tái sử dụng, quản lý state hiệu quả, và cập nhật UI động, mang lại trải nghiệm người dùng mượt mà.

2.3.3. Giao diện người dùng

Giao diện người dùng được thiết kế tập trung vào trải nghiệm người dùng, đảm bảo tính trực quan, thân thiện, dễ sử dụng và cung cấp đầy đủ thông tin hỗ trợ:

Khởi tải lên file Excel: Cho phép người dùng tải lên file dữ liệu đầu vào. Giao diện hiển thị hướng dẫn rõ ràng về định dạng file, cung cấp mẫu file Excel để người dùng tham khảo và liệt kê các thuộc tính dữ liệu cần thiết. Hệ thống cũng tích hợp cơ chế xử lý lỗi để thông báo cho người dùng về các vấn đề tiềm ẩn, chẳng hạn như định dạng file không hợp lệ, dữ liệu thiếu hoặc dữ liệu không đúng định dạng.

Khởi hiển thị tiến trình: Hiển thị trực quan tiến trình xử lý dữ liệu và thực thi giải thuật, cho phép người dùng theo dõi trạng thái và ước tính thời gian hoàn thành. Thông tin tiến trình được cập nhật theo thời gian thực, giúp người dùng nắm bắt được trạng thái hiện tại của hệ thống.

Khởi hiển thị kết quả: Trình bày kết quả ghép cặp dưới dạng bảng, cho phép người dùng xem chi tiết từng cặp sinh viên - cố vấn, bao gồm thông tin của cả sinh viên và cố vấn. Giao diện cung cấp các chức năng sắp xếp và lọc dữ liệu, giúp người dùng dễ dàng tìm kiếm và phân tích kết quả theo nhu cầu.

Khởi hiển thị phân tích chuyên sâu (insights): Trực quan hóa dữ liệu phân tích, sử dụng biểu đồ và đồ thị tương tác giúp người dùng dễ dàng nắm bắt thông tin và khám phá dữ liệu chi tiết hơn. Người dùng có thể tương tác với biểu đồ để lọc, sắp xếp và xem chi tiết dữ liệu.

CHƯƠNG 3 – CÀI ĐẶT VÀ THỬ NGHIỆM

3.1. Thiết kế thực nghiệm

3.1.1. Môi trường thực nghiệm

Hệ hỗ trợ quyết định ghép cặp sinh viên – cố vấn học tập được xây dựng trong đề án được triển khai trên máy chủ với cấu hình như sau:

Hệ điều hành: Microsoft Windows 10

CPU: Intel(R) Core(TM) i5-13400F thế hệ thứ 13 (10 nhân, 16 luồng)

RAM: 32 GB

Cấu hình này đảm bảo hệ hỗ trợ quyết định có đủ tài nguyên để xử lý các tập dữ liệu lớn và thực thi các giải thuật phức tạp trong thời gian hợp lý.

3.1.2. Bộ dữ liệu thực nghiệm

Do tính chất riêng tư và bảo mật của dữ liệu sinh viên và cố vấn, việc thu thập dữ liệu thực tế gặp nhiều khó khăn. Vì vậy, dữ liệu được sử dụng trong nghiên cứu này được mô phỏng dựa trên dữ liệu thực tế và được tổng hợp, điều chỉnh để phù hợp với mục tiêu nghiên cứu và đảm bảo đủ số lượng mẫu cho việc đánh giá giải thuật. Bộ dữ liệu đầy đủ bao gồm thông tin của 500 sinh viên và 30 cố vấn, đi kèm 10 thuộc tính cho mỗi sinh viên và cố vấn, mô tả chi tiết các đặc điểm, năng lực, nhu cầu và sở thích của từng đối tượng. Mỗi sinh viên và cố vấn sẽ tự xác định nhu cầu và sở thích của mình, đồng thời gán trọng số cho các thuộc tính của đối phương, thể hiện mức độ ưu tiên của họ cho từng tiêu chí trong quá trình ghép cặp.

3.1.3. Kịch bản thử nghiệm

Kịch bản thử nghiệm được triển khai là sử dụng một tập dữ liệu lớn với số lượng lớn sinh viên và cố vấn, gần với thực tế. Kịch bản này được thiết kế để đánh giá khả năng mở rộng của giải thuật, tức là khả năng xử lý một lượng lớn dữ liệu và nhiều thuộc tính. Thử nghiệm này cũng giúp đánh giá hiệu suất tính toán của giải thuật khi không gian tìm kiếm trở nên phức tạp hơn.

3.1.4. Tiêu chí đánh giá

Để so sánh và đánh giá hiệu quả của các giải thuật, hai tiêu chí đánh giá chính được sử dụng:

Giá trị hàm fitness: Hàm fitness đo lường chất lượng của một giải pháp (một phương án ghép cặp). Trong bài toán ghép cặp sinh viên - cố vấn, hàm fitness được thiết kế để phản ánh mức độ phù hợp và hài lòng của cả sinh viên và cố vấn trong cặp ghép đó. Giá trị fitness được tính toán dựa trên sự tương đồng giữa các thuộc tính của sinh viên và cố vấn (chuyên ngành, kinh nghiệm, phong cách nghiên cứu,...) cũng như trọng số mà mỗi bên gán cho các thuộc tính đó. Giá trị fitness càng cao thể hiện chất lượng cặp ghép càng tốt, tức là mức độ hài lòng của cả sinh viên và cố vấn càng cao. Phân tích giá trị fitness cho phép chúng ta đánh giá khả năng của giải thuật trong việc tìm ra các cặp ghép tối ưu.

Thời gian chạy: Thời gian chạy của giải thuật là thời gian cần thiết để giải thuật tìm ra giải pháp tối ưu (hoặc một tập hợp các giải pháp tối ưu Pareto trong trường hợp tối ưu đa mục tiêu). Thời gian chạy càng ngắn, giải thuật càng hiệu quả về mặt tính toán. Tiêu chí này đặc

biệt quan trọng khi xử lý các tập dữ liệu lớn, vì thời gian chạy có thể tăng đáng kể theo kích thước của bài toán. So sánh thời gian chạy giữa các giải thuật khác nhau giúp xác định giải thuật nào có hiệu suất tính toán tốt nhất cho bài toán ghép cặp sinh viên - cố vấn.

Cụ thể hơn, để so sánh hiệu năng cuối cùng của các giải thuật khác nhau bằng một giá trị số duy nhất, một hàm fitness tổng hợp (aggregate fitness function) được sử dụng. Hàm này được áp dụng sau khi giải thuật đã chạy xong, nhằm mục đích lượng hóa chất lượng tổng thể của một giải pháp. Với việc giá trị fitness được tính toán dựa trên sự tương đồng giữa các thuộc tính của sinh viên và cố vấn, một công thức hàm fitness tổng hợp phù hợp có thể được định nghĩa là một tổng có trọng số của các điểm tương thích.

Việc sử dụng hàm fitness tổng hợp như trên là hợp lý để đánh giá hiệu quả của mô hình đề xuất. Thứ nhất, hàm này có tính đại diện cao vì nó trực tiếp lượng hóa các mục tiêu cốt lõi của bài toán. Về bản chất, nó là một dạng cụ thể hóa của hàm mục tiêu f_4 (tối đa hóa sự phù hợp tổng thể), vốn được thiết kế để đo lường chất lượng song phương của các cặp ghép. Một giá trị fitness cao cho thấy giải pháp đã thành công trong việc tạo ra các cặp ghép có độ tương thích cao từ cả hai phía, qua đó đáp ứng tốt các mục tiêu f_1 (tối đa hóa sự hài lòng của sinh viên) và f_2 (tối đa hóa sự hài lòng của cố vấn). Thứ hai, quá trình tối ưu hóa đồng thời các hàm mục tiêu f_1 , f_2 , f_4 , f_5 và cân bằng f_3 một cách tự nhiên sẽ hướng đến việc tìm kiếm các giải pháp có giá trị fitness tổng hợp cao. Cuối cùng, việc sử dụng một hàm đánh giá tổng hợp, nhất quán để so sánh kết quả cuối cùng của các giải thuật là một phương pháp khoa học tiêu chuẩn, cung cấp một thước đo chung khách quan để khẳng định mô hình đề xuất hoạt động tốt và hiệu quả.

3.2. Cài đặt hệ thống và kết quả thực nghiệm

3.2.1. Cài đặt

Hệ hỗ trợ quyết định được xây dựng trong đề án được phát triển bằng Java 17, Spring Boot và MOEA Framework cho backend và React cho frontend. Backend sử dụng MOEA Framework để triển khai các giải thuật tối ưu tiến hóa đa mục tiêu. Các tham số của giải thuật được cấu hình như sau: số lần lặp (10), số thế hệ lai ghép (100), kích thước quần thể (1000) và giới hạn thời gian tối ưu hóa (5000 milliseconds).

3.2.2. Kết quả thực nghiệm

Kết quả thực nghiệm cho thấy giải thuật NSGA-III, khi được áp dụng cho bài toán ghép cặp sinh viên - cố vấn học tập, đã đạt được giá trị hàm fitness là 31470.809 và thời gian chạy là 6052.000 ms (tương đương 6,052 giây). Giá trị fitness này phản ánh chất lượng của phương án ghép cặp được tạo ra bởi giải thuật. Một giá trị fitness cao hơn thường cho thấy mức độ phù hợp giữa sinh viên và cố vấn tốt hơn. Thời gian chạy 6,052 giây cho thấy hiệu suất tính toán của NSGA-III trên bộ dữ liệu được sử dụng trong thực nghiệm. Thời gian này là tương đối nhanh khi xét vào quy mô của bài toán (số lượng sinh viên và cố vấn) và cấu hình máy tính được sử dụng.

3.3. Đánh giá

3.3.1. So sánh với các phương pháp khác

Trong quá trình triển khai thực nghiệm, ngoài việc sử dụng giải thuật NSGA-III, năm giải thuật khác gồm eMOEA, VEGA, NSGA-II, IBEA, PESA2 cũng được triển khai sử dụng với mục đích so sánh với giải thuật NSGA-III về giá trị fitness và thời gian chạy, từ đó phân tích ưu điểm và vấn đề còn tồn đọng của giải thuật đề xuất.

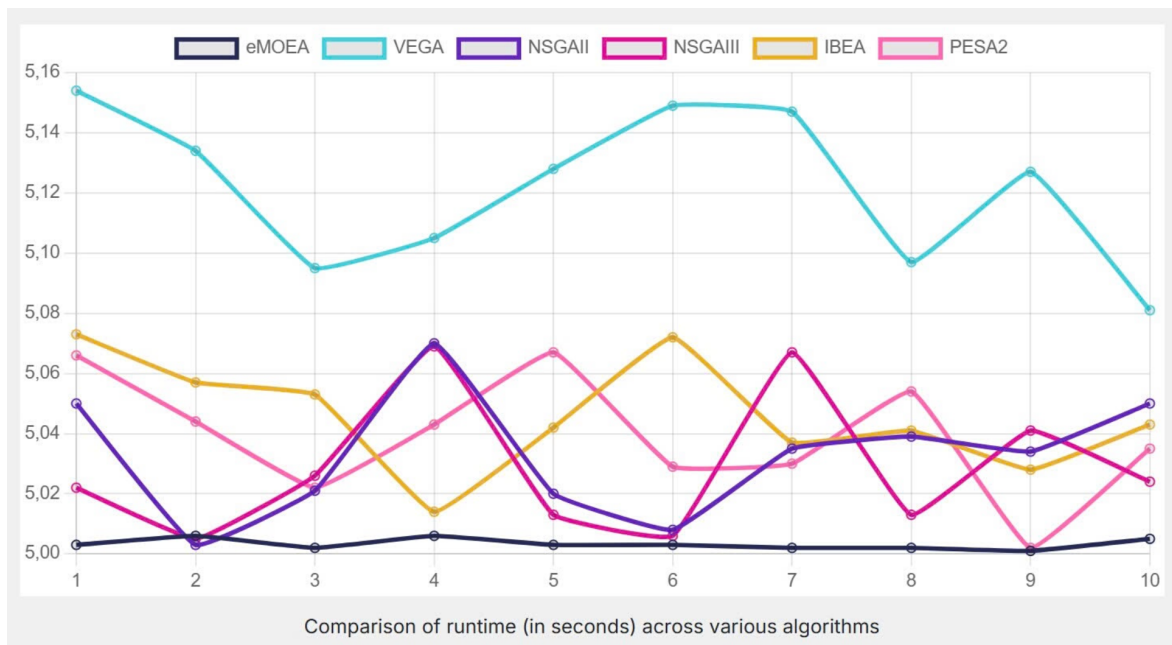
Bảng 3.1: Bảng so sánh kết quả giá trị fitness giữa các giải thuật

Iteration	eMOEA	VEGA	NSGAI	NSGAIII	IBEA	PESA2
1	31487,73	31469,83	31503,98	31517,44	31511,83	31487,14
2	31494,23	31490,6	31510,07	31515,07	31523,14	31500,12
3	31494,32	31474,63	31506,62	31514,4	31508,65	31496,33
4	31486,28	31473,62	31514,34	31511,11	31506,62	31486,13
5	31492,65	31485,96	31496,32	31508,28	31507,96	31503,33
6	31493,11	31490,98	31507,81	31512,86	31494,95	31472
7	31497,57	31477,91	31507,12	31521,53	31509,63	31484,69
8	31493,47	31481,16	31506,35	31506,35	31520,57	31492,89
9	31482,15	31486,47	31514,01	31504,2	31505,96	31492,78
10	31498,79	31485,28	31501,12	31512,94	31501,63	31503,13

Bảng 3.2: Bảng so sánh kết quả thời gian chạy của các giải thuật

Iteration	eMOEA	VEGA	NSGAI	NSGAIII	IBEA	PESA2
-----------	-------	------	-------	---------	------	-------

1	5,003	5,154	5,05	5,022	5,073	5,066
2	5,006	5,134	5,003	5,005	5,057	5,044
3	5,002	5,095	5,021	5,026	5,053	5,022
4	5,006	5,105	5,07	5,069	5,014	5,043
5	5,003	5,128	5,02	5,013	5,042	5,067
6	5,003	5,149	5,008	5,006	5,072	5,029
7	5,002	5,147	5,035	5,067	5,037	5,03
8	5,002	5,097	5,039	5,013	5,041	5,054
9	5,001	5,127	5,034	5,041	5,028	5,002
10	5,005	5,081	5,05	5,024	5,043	5,035



Hình 3.1: Đồ thị biểu diễn kết quả thời gian chạy của các giải thuật

Từ dữ liệu trong Bảng 3.1, có thể rút ra những phân tích sâu sắc về hiệu năng của từng giải thuật. NSGA-III thể hiện một hiệu năng vượt trội và ổn định nhất về mặt chất lượng giải pháp. Trong 10 lần chạy lặp, NSGA-III đạt được giá trị fitness value cao nhất trong 9 lần (từ iteration 1 đến iteration 8 và iteration 10). Duy nhất ở iteration 9, giá trị của NSGA-III (31504,20) thấp hơn một chút so với NSGA-II (31514,01) và IBEA (31505,96), nhưng vẫn cao hơn các giải thuật còn lại.

Khi xét về giá trị trung bình, ưu điểm của NSGA-III càng trở nên rõ ràng. Với giá trị fitness value trung bình là 31512,42, NSGA-III đứng đầu một cách thuyết phục. Giải thuật có

hiệu năng gần nhất là IBEA với giá trị trung bình là 31509,09, theo sau là NSGA-II với 31506,77. Điều này cho thấy NSGA-III không chỉ thường xuyên tìm ra các giải pháp tốt nhất trong từng lần chạy riêng lẻ mà còn duy trì được một mức hiệu năng trung bình cao và ổn định. Sự khác biệt giữa NSGA-III và nhóm các giải thuật có hiệu năng thấp hơn như VEGA (trung bình 31481,64) là rất đáng kể, cho thấy sự ưu việt của cơ chế tối ưu hóa trong NSGA-III.

Sự vượt trội của NSGA-III có thể được lý giải bởi cơ chế lựa chọn dựa trên điểm tham chiếu, vốn được thiết kế đặc biệt để xử lý hiệu quả các bài toán có nhiều mục tiêu (many-objective problems). Cơ chế này giúp duy trì sự đa dạng của các giải pháp trên mặt trận Pareto một cách tốt hơn so với cơ chế đo lường mật độ của NSGA-II, đặc biệt khi số lượng mục tiêu tăng lên. Trong khi IBEA cũng cho thấy hiệu năng khá tốt, sự ổn định và hiệu năng trung bình cao hơn của NSGA-III đã khẳng định vị thế của nó là lựa chọn hàng đầu trong bối cảnh thực nghiệm này.

Phân tích từ bảng kết quả thu được về thời gian chạy của các giải thuật cho thấy một bức tranh khác biệt so với phân tích về chất lượng giải pháp. Giải thuật eMOEA thể hiện ưu thế vượt trội về tốc độ tính toán, với thời gian chạy trung bình chỉ là 5,003 giây, nhanh nhất trong số các giải thuật được so sánh. Tuy nhiên, khi nhìn vào hiệu năng của NSGA-III, một điểm rất đáng chú ý là nó có thời gian chạy trung bình là 5,029 giây, đứng thứ hai về tốc độ và chỉ chậm hơn một chút không đáng kể so với eMOEA. Điều này đặc biệt ấn tượng khi đặt trong bối cảnh chất lượng giải pháp vượt trội của nó. NSGA-III nhanh hơn đáng kể so với VEGA (5,122 giây) và nhanh hơn một cách nhất quán so với các giải thuật còn lại là NSGA-II (5,038 giây), PESA2 (5,039 giây), và IBEA (5,046 giây). Điều này bác bỏ giả định rằng một giải thuật phức tạp hơn về mặt khái niệm như NSGA-III sẽ luôn chậm hơn các đối thủ. Có thể cơ chế lựa chọn dựa trên điểm tham chiếu của NSGA-III lại có thể hiệu quả hơn về mặt tính toán so với các cơ chế khác.

3.3.2. Nhận xét, ưu điểm, hạn chế

Giải pháp đề xuất có nhiều ưu điểm nổi bật, trong đó quan trọng nhất là khả năng tìm kiếm và cung cấp một tập hợp đa dạng các giải pháp tối ưu (mặt trận Pareto). Điều này cho phép người ra quyết định có nhiều lựa chọn linh hoạt, thay vì bị giới hạn trong một phương án duy nhất. Hơn nữa, hệ thống đảm bảo tính bền vững cao cho các cặp ghép bằng cách tích hợp các nguyên lý của lý thuyết ghép cặp bền vững, giúp tạo ra các mối quan hệ ổn định và

lâu dài. Giải pháp này cũng thể hiện sự vượt trội so với các phương pháp truyền thống nhờ tính khách quan, linh hoạt, khả năng tùy chỉnh cao và xem xét được đồng thời nhiều yếu tố phức tạp.

Tuy nhiên, giải pháp cũng tồn tại một số hạn chế cần nhìn nhận. Độ phức tạp tính toán có thể trở thành một thách thức đối với các bài toán quy mô rất lớn, và hiệu quả của hệ thống phụ thuộc chặt chẽ vào chất lượng và sự đầy đủ của dữ liệu đầu vào. Thêm vào đó, việc hệ thống cung cấp một tập hợp các giải pháp tối ưu tuy là một ưu điểm nhưng cũng tạo ra thách thức mới trong việc lựa chọn một giải pháp cuối cùng để triển khai. Cuối cùng, việc lượng hóa các yếu tố định tính phức tạp như sự tương thích về tính cách vẫn còn hạn chế và mô hình hiện tại chủ yếu xử lý việc ghép cặp theo lô, chưa linh hoạt với các thay đổi động của hệ thống.

KẾT LUẬN

Đề án đã tập trung vào việc giải quyết bài toán ghép cặp sinh viên - cố vấn học tập, một vấn đề quan trọng trong việc tối ưu hóa hiệu quả hoạt động giáo dục. Việc ghép cặp hiệu quả không chỉ giúp sinh viên tiếp cận được sự hướng dẫn phù hợp với nhu cầu và mục tiêu học tập cá nhân, mà còn hỗ trợ các cố vấn trong việc phân bổ thời gian và nguồn lực một cách hiệu quả. Đề án đồng thời cũng đạt được những kết quả đáng kể trong việc tiếp cận bài toán này một cách khoa học và hệ thống, đóng góp vào việc cải thiện quy trình cố vấn học tập.

Đầu tiên, một mô hình toán học chính quy đã được xây dựng để mô tả bài toán ghép cặp sinh viên - cố vấn. Mô hình này không chỉ xác định rõ ràng tập hợp sinh viên và cố vấn tham gia vào quá trình ghép cặp, mà còn thiết lập một hệ thống tiêu chí đánh giá sự phù hợp giữa hai bên. Hàm mục tiêu của mô hình được thiết kế để tối ưu hóa sự hài lòng chung của cả sinh viên và cố vấn, đồng thời các ràng buộc được đưa ra để đảm bảo tính khả thi và thực tiễn của việc ghép cặp.

Thứ hai, một giải thuật lai hiệu quả đã được phát triển, kết hợp sức mạnh của lý thuyết ghép cặp bền vững (Stable matching theory) và giải thuật NSGA-III. Lý thuyết ghép cặp bền vững đảm bảo tính ổn định của kết quả, trong khi NSGA-III cho phép tối ưu hóa đồng thời nhiều mục tiêu. Sự kết hợp này tạo ra một giải thuật lai mạnh mẽ, vừa đảm bảo tính ổn định, vừa linh hoạt trong việc đáp ứng các yêu cầu đa dạng của bài toán.

Cuối cùng, đề án đã phát triển một hệ thống hỗ trợ quyết định với giao diện người dùng thân thiện, đóng vai trò là cầu nối giữa mô hình toán học, giải thuật tối ưu và người dùng cuối, giúp đơn giản hóa quy trình và khuyến khích sự tham gia tích cực của các bên.

Đề xuất cốt lõi của đề án - sử dụng giải thuật lai SMT-NSGA-III - đã được chứng minh là có hiệu quả vượt trội thông qua các thử nghiệm thực nghiệm. Phân tích so sánh cho thấy NSGA-III không chỉ tạo ra các giải pháp có chất lượng cao hơn đáng kể (giá trị fitness cao hơn) so với các giải thuật MOEA khác như NSGA-II, IBEA và PESA2, mà còn duy trì được hiệu suất tính toán rất cạnh tranh. Điều này khẳng định rằng phương pháp được chọn mang lại sự cân bằng tối ưu giữa chất lượng giải pháp và hiệu quả thực thi.

Đóng góp quan trọng nhất của đề án là việc tạo ra một hệ thống có khả năng đưa ra các phương án ghép cặp vừa tối ưu đa mục tiêu, vừa bền vững. Giải pháp này vượt xa các phương pháp ghép cặp truyền thống, cung cấp một công cụ khoa học, khách quan và hiệu quả để hỗ trợ các nhà quản lý giáo dục, góp phần trực tiếp vào việc nâng cao chất lượng giáo dục, tối ưu hóa nguồn lực và tăng cường sự hài lòng cho cả sinh viên và cô vấn.