



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0034619

(51)⁷ G06F 17/50

(13) B

(21) 1-2019-01416

(22) 21/03/2019

(45) 25/01/2023 418

(43) 25/07/2019 376A

(73) Tập đoàn Công nghiệp - Viễn thông Quân đội (VN)

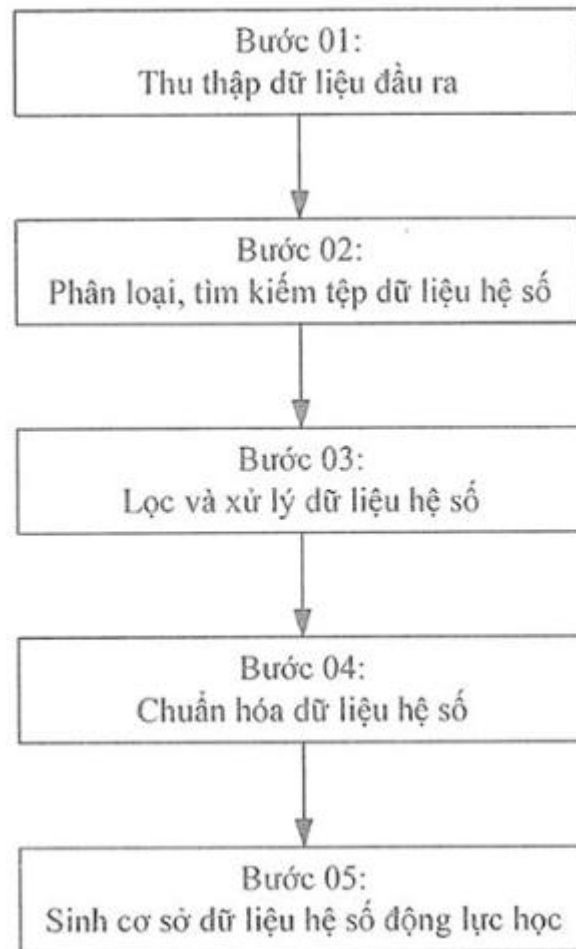
Số 1 đường Trần Hữu Dực, phường Mỹ Đình 2, quận Nam Từ Liêm, thành phố Hà Nội

(72) Nguyễn Hải Anh (VN); Nguyễn Duy Trọng (VN).

(74) Công ty TNHH Tư vấn Quốc Dân (NACILAW)

(54) PHƯƠNG PHÁP VÀ HỆ THỐNG TỰ ĐỘNG TÌM KIẾM, XỬ LÝ VÀ TỔNG HỢP DỮ LIỆU LỚN TRONG BÀI TOÁN HẬU XỬ LÝ TÍNH TOÁN ĐỘNG LỰC HỌC CHẤT LỎNG

(57) Sáng chế đề cập đến phương pháp và hệ thống tự động tìm kiếm, phân tích và tổng hợp dữ liệu lớn phục vụ bài toán hậu xử lý tính toán động lực học chất lỏng. Phương pháp bao gồm năm bước: bước 1: thu thập dữ liệu đầu ra từ chương trình tính toán chất lỏng; bước 2: phân loại và tìm kiếm dữ liệu; bước 3: lọc và xử lý dữ liệu; bước 4: chuẩn hóa dữ liệu hệ số; bước 5: sinh cơ sở dữ liệu hệ số động lực học. Hệ thống tự động tìm kiếm, phân tích và tổng hợp dữ liệu lớn được xây dựng dựa vào phương pháp trên bao gồm ba mô đun: mô đun tìm kiếm dữ liệu; mô đun phân tích dữ liệu, mô đun phân phối dữ liệu.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp và hệ thống tự động tìm kiếm, xử lý và tổng hợp dữ liệu phục vụ bài toán hậu xử lý tính toán động lực học chất lỏng với số lượng trường hợp tính toán lớn. Sáng chế được ứng dụng trong quá trình xây dựng cơ sở dữ liệu động lực học chất lỏng phục vụ bài toán mô hình hóa, mô phỏng chuyển động các vật thể hoạt động trong môi trường không khí, nước hoặc không khí và nước.

Tình trạng kỹ thuật của sáng chế

Phương pháp và hệ thống tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trình bày cách thức và công cụ để chuyển dữ liệu từ dạng thô, phân tán (đầu ra của chương trình tính toán chuyên dụng) trở thành dạng tinh và tập trung (cơ sở dữ liệu hệ số động lực học chất lỏng). Thông thường, việc xây dựng cơ sở dữ liệu hệ số động lực học được thực hiện thủ công, tuân tự cho từng trường hợp tính toán riêng biệt và số lượng trường hợp không lớn. Theo phương pháp này, người thực hiện xây dựng cơ sở dữ liệu hệ số cần thực hiện lặp đi lặp lại các bước sau đây cho mỗi một trường hợp tính toán động lực học: bước 1: xác định vị trí lưu trữ toàn bộ đầu ra của chương trình tính toán; bước 2: xác định vị trí lưu trữ các tệp dữ liệu hệ số; bước 3: kiểm tra số lượng và chủng loại các hệ số; bước 4: kiểm tra tính hội tụ của dữ liệu hệ số; bước 5: tính toán giá trị đại diện; bước 6: sắp xếp giá trị đại diện vào cơ sở dữ liệu hệ số. Phương pháp xử lý dữ liệu truyền thống này được minh họa chi tiết trong Hình 1.

Nhược điểm của phương pháp truyền thống là việc tính toán và xử lý hầu hết được thực hiện thủ công nên hiệu suất không cao và không thể thực hiện trên số lượng dữ liệu lớn như yêu cầu của bài toán xây dựng cơ sở dữ liệu hệ số. Ngoài ra, phương pháp truyền thống không có khả năng đưa ra các khuyến cáo hoặc thông báo lỗi về cấu trúc dữ liệu hoặc kết quả đầu ra của chương trình tính toán.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là đề xuất phương pháp và hệ thống tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng nhằm sinh cơ sở dữ liệu hệ số dùng trong mô phỏng chuyển động sáu bậc tự do của các loại vật thể. Trong đó, cơ sở dữ liệu hệ số động lực học chất lỏng đặc trưng cho đáp ứng vật lý của vật thể tại các trạng thái khác nhau trong miền hoạt động. Bên cạnh đó, trong quá trình triển khai, phương pháp và hệ thống đề xuất có sử dụng thuật toán tiêu chuẩn hội tụ và các thuật toán kiểm tra cấu trúc dữ liệu để đưa ra các thông báo khi cần thiết.

Để đạt được mục đích nêu trên, phương pháp tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng cần thực hiện năm bước như sau: bước 1: thu thập dữ liệu đầu ra từ chương trình tính toán chất lỏng; bước 2: phân loại, tìm kiếm các tệp chứa hệ số; bước 3: lọc, xử lý dữ liệu hệ số; bước 4: chuẩn hóa dữ liệu hệ số; bước 5: sinh cơ sở dữ liệu hệ số. Hệ thống tự động tìm kiếm, xử lý và tổng hợp dữ liệu được triển khai dựa trên phương pháp trên. Hệ thống bao gồm ba môđun như sau: môđun tìm kiếm dữ liệu; môđun xử lý dữ liệu; môđun phân phối dữ liệu.

Cụ thể, chi tiết các bước trong phương pháp tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn được trình bày trong các phần tiếp theo.

Bước 1: thu thập dữ liệu đầu ra từ chương trình tính toán động lực học chất lỏng. Đầu ra của một trường hợp tính toán động lực học chất lỏng là một tệp chứa tất cả các thông tin quan trọng như hình học, lưới rời rạc, thiết lập tính toán, kết quả tính toán... Để xây dựng cơ sở dữ liệu hệ số, cần sử dụng rất nhiều trường hợp tính toán. Do đó, số lượng tệp đầu ra từ chương trình tính toán là cực kỳ lớn và thường được lưu trữ phân tán, không có thứ tự. Vì vậy, mục đích của bước đầu tiên là thu thập tất cả các tệp thông tin và sắp xếp vào một tệp duy nhất. Như vậy, đầu vào của bước này là các tệp thông tin của các trường hợp tính toán bị lưu trữ phân tán, đầu ra là một tệp tin duy nhất chứa tất cả các tệp thông tin tính toán được sắp xếp theo thứ tự thời gian phục vụ quá trình xử lý sau này. Việc tổng hợp này có thể được thực hiện trực tiếp hoặc gián tiếp. Thu thập trực tiếp là

chuyển các tệp thông tin tính toán về tệp tin tổng hợp. Thu thập gián tiếp là chỉ thu thập địa chỉ lưu trữ các tệp thông tin tính toán.

Bước 2: phân loại và tìm kiếm tệp dữ liệu hệ số. Mục đích đầu tiên của bước này là phân loại các thông tin trong tệp đầu ra theo quy ước đặt tên có sẵn sinh bởi chương trình tính toán. Ví dụ: trong tệp dữ liệu của trường hợp thứ nhất, sẽ có các tệp con chứa thông tin về hình học (geometry), lưới rời rạc (mesh)... Trong đó, tệp được quan tâm trong bước này là tệp chứa các hệ số đầu ra, được nhận dạng thông qua tên của tệp. Sau đó, mục đích thứ hai của bước này là tìm kiếm tệp con chứa loại hệ số theo yêu cầu từ người sử dụng trong tệp chứa các hệ số đầu ra. Thông thường, chương trình tính toán sinh sáu tệp chứa hệ số bao gồm: Cx, hệ số lực dọc thân; Cy, hệ số lực trượt cạnh; Cz, hệ số lực pháp tuyến; Cl, hệ số mô men quanh trục dọc; Cm, hệ số mô men chúc góc; Cn, hệ số mô men trượt cạnh. Mục đích thứ hai của bước 2 nhằm tìm kiếm loại hệ số mà người sử dụng mong muốn tổng hợp bằng phương pháp kiểm trùng tên. Như vậy, đầu vào của bước 2 là các tệp chứa dữ liệu đầu ra đã được thu thập trong bước 1, đầu ra là các tệp con chứa loại hệ số theo yêu cầu của người sử dụng. Ngoài ra, trong bước này, phương pháp đề xuất thêm thuật toán kiểm tra về số lượng và chủng loại của các tệp chứa hệ số. Nếu đầu ra của chương trình tính toán không đủ số lượng là 6 và chủng loại (Cx, Cy, Cz, Cl, Cm, Cn) thì sẽ đưa ra thông báo tới người sử dụng. Việc phân loại và tìm kiếm dữ liệu được thực hiện tự động bằng thuật toán nhận dạng tên của tệp trong môđun tìm kiếm dữ liệu của hệ thống đề xuất.

Bước 3: lọc và xử lý dữ liệu hệ số. Tệp chứa dữ liệu của một loại hệ số thường là một danh sách các số sinh bởi chương trình tính toán theo thời gian. Các số này thường biểu diễn một dao động dạng hình sin. Trong bước này, tệp chứa dữ liệu hệ số từ bước 2 được kiểm tra về tính hội tụ sử dụng thuật toán hội tụ trong môđun xử lý dữ liệu của hệ thống đề xuất. Cụ thể, dữ liệu hệ số được coi là hội tụ nếu thỏa mãn điều kiện sau: thứ nhất là biên độ dao động của năm chu kỳ cuối cùng giảm nhiều hơn 10%; thứ hai là biên độ dao động của năm chu kỳ cuối không quá 10% giá trị cực tiểu. Các tệp dữ liệu hệ số không hội tụ sẽ được lọc bỏ và gửi thông báo đến người dùng. Thông báo này có hai cấp độ chính: thứ nhất là dữ liệu không hội tụ (không thỏa mãn điều kiện thứ nhất và thứ hai);

thứ hai là dữ liệu cần tính toán thêm (thỏa mãn một trong hai điều kiện). Ngoài ra, các tiêu chuẩn có thể được điều chỉnh bởi người dùng tùy theo yêu cầu thực tế của bài toán. Khi đó, đầu vào của bước 3 là các tệp chứa dữ liệu hệ số theo yêu cầu của người dùng và đầu ra sau bước lọc chỉ còn lại các tệp chứa dữ liệu hội tụ theo tiêu chuẩn hội tụ. Mục đích của bước xử lý dữ liệu hệ số là tính toán giá trị trung bình cho toàn bộ danh sách số trong tệp dữ liệu hệ số. Giá trị này là giá trị trung bình của khoảng năm chu kỳ cuối trong dữ liệu hội tụ và được gọi là giá trị đại diện. Tương tự, việc đọc và tính toán giá trị đại diện cũng được thực hiện trong môđun xử lý dữ liệu hệ số. Như vậy, đầu ra cuối cùng của bước 3 là các giá trị hệ số đại diện. Mỗi một giá trị đó đại diện cho toàn bộ nội dung bên trong tệp dữ liệu hệ số tương ứng.

Bước 4: chuẩn hóa dữ liệu hệ số. Trong bước này, giá trị đại diện là đầu ra của bước 3 sẽ được chuẩn hóa về định dạng theo tiêu chuẩn của cơ sở dữ liệu hệ số. Cụ thể, giá trị đại diện được đưa về dạng số thực có 4 chữ số thập phân sau dấu phẩy. Bên cạnh đó, giá trị đại diện được kiểm tra so sánh với giá trị cực đại và giá trị cực tiểu của loại hệ số tương ứng. Hơn nữa, hệ số động lực học chất lỏng thường có quy luật nhất định, nên giá trị đại diện cũng được kiểm tra đối chiếu với các giá trị đại diện lân cận. Ví dụ: hệ số lực pháp tuyến C_z thường tăng tuyến tính tại một giới hạn góc tấn. Do đó, nếu giá trị đại diện không tuân theo quy luật trên thì khả năng đầu ra của chương trình tính toán là sai. Như vậy, đầu vào của bước chuẩn hóa dữ liệu là giá trị hệ số đại diện được xác định trong bước 3, đầu ra là giá trị đã được chuẩn hóa thỏa mãn các yêu cầu sau: thứ nhất là định dạng chuẩn; thứ hai là nằm trong giới hạn cho phép; thứ ba là phù hợp với quy luật biến thiên của hệ số. Nếu giá trị đại diện không thỏa mãn một trong các điều kiện trên, thông báo lỗi sẽ được đưa ra tương ứng. Bước chuẩn hóa dữ liệu được thực hiện nhờ môđun xử lý dữ liệu trong hệ thống đề xuất.

Bước 5: sinh cơ sở dữ liệu hệ số động lực học. Mục đích của bước này là xây dựng cơ sở dữ liệu loại hệ số theo yêu cầu của người dùng từ các giá trị đại diện đã chuẩn hóa. Cơ sở dữ liệu thường là các bảng dạng hai chiều hoặc ba chiều được sắp xếp theo một thứ tự nhất định. Mỗi một giá trị trong cơ sở dữ liệu tương ứng với một trường hợp từ chương trình tính toán động lực học chất lỏng. Do đó, giá trị đại diện từ bước 4 được phân phối

vào cơ sở dữ liệu theo thứ tự tương ứng với thứ tự được sắp xếp từ bước 1. Đầu ra của bước 5 là bảng dữ liệu chứa các hệ số đại diện được tính trong bước 4 theo một thứ tự được sắp xếp từ trước. Nếu cơ sở dữ liệu chưa tồn tại thì nó sẽ được tự động khởi tạo theo định dạng của loại hệ số (hai chiều hoặc ba chiều) nhờ môđun phân phối dữ liệu trong hệ thống đề xuất. Ngoài ra, nếu vị trí trong cơ sở dữ liệu đã có giá trị khác đang được lưu trữ, môđun phân phối sẽ đưa ra thông báo và cho phép người dùng chọn giữ nguyên giá trị cũ hoặc thay thế bằng giá trị mới.

Để triển khai phương pháp trên, hệ thống tự động tìm kiếm, xử lý và tổng hợp dữ liệu được xây dựng để hỗ trợ quá trình hậu xử lý tính toán động lực học chất lỏng. Cụ thể, hệ thống được tổ chức thành ba môđun chính: môđun tìm kiếm dữ liệu, môđun xử lý dữ liệu và môđun phân phối dữ liệu. Trong đó, môđun tìm kiếm dữ liệu cho phép tự động thực hiện bước 2 nhằm xác định chính xác tệp chứa dữ liệu hệ số cần xử lý. Sau đó, môđun xử lý dữ liệu thực hiện bước 3 và bước 4 để đưa ra giá trị hệ số động lực học chất lỏng đại diện đã được chuẩn hóa. Cuối cùng, môđun phân phối dữ liệu có chức năng thực hiện bước 5 sinh cơ sở dữ liệu để lưu trữ hệ số động lực học chất lỏng phục vụ mô phỏng chuyển động của các loại vật thể.

Mô tả vắn tắt các hình vẽ

Hình 1 là hình vẽ minh họa sơ đồ tổng hợp, phân tích và tổng hợp dữ liệu cơ sở động lực học theo phương pháp truyền thống;

Hình 2 là hình vẽ minh họa sơ đồ tổng hợp, phân tích và phân phối dữ liệu hệ số động lực học sử dụng phương pháp mới;

Hình 3 là hình vẽ minh họa các bước thực hiện tổng hợp, phân tích và phân phối dữ liệu hệ số vào cơ sở dữ liệu;

Hình 4 là hình vẽ minh họa thuật toán nhận dạng đồ thị hệ số;

Hình 5 là hình vẽ minh họa sơ đồ thuật toán môđun tìm kiếm dữ liệu hệ số động lực học chất lỏng;

Hình 6 là hình vẽ minh họa sơ đồ thuật toán môđun xử lý dữ liệu;

Hình 7 là hình vẽ minh họa sơ đồ thuật toán môđun phân phối dữ liệu hệ số động lực học;

Hình 8 là hình vẽ minh họa cấu trúc tệp đầu ra của chương trình tính toán.

Mô tả chi tiết sáng chế

Tham chiếu đến Hình 1 và Hình 2, sáng chế đề cập đến phương pháp và hệ thống xử lý dữ liệu số lượng lớn nhằm quy chuẩn hóa và tự động hóa quá trình tìm kiếm, phân tích và tổng hợp dữ liệu phục vụ bài toán hậu xử lý tính toán hệ số động lực học chất lỏng và xây dựng cơ sở dữ liệu trong mô hình hóa, mô phỏng chuyển động sáu bậc tự do của các loại vật thể như máy bay, tàu thủy, ô tô... Phương pháp và hệ thống mới được đưa ra nhằm giải quyết những hạn chế của phương pháp truyền thống và nâng cao hiệu quả tổng hợp và xử lý dữ liệu, đặc biệt với số lượng dữ liệu lớn như việc xây dựng cơ sở dữ liệu các bài toán mô phỏng trong thực tế.

Trước tiên, cơ sở dữ liệu hệ số động lực học chất lỏng được hiểu là các bảng hai chiều hoặc ba chiều lưu trữ các hệ số đặc trưng cho đáp ứng của vật thể tại tất cả các trạng thái hoạt động. Cơ sở dữ liệu hệ số bao gồm sáu loại hệ số chính bao gồm: Cx – đặc trưng cho tác dụng lực dọc thân vật thể; Cy – đặc trưng cho tác dụng lực trượt cạnh lên vật thể; Cz – đặc trưng cho tác dụng lực pháp tuyến với thân vật thể; Cl – đặc trưng cho mô men tác dụng quanh trục dọc thân; Cm – đặc trưng cho mô men chúc góc; Cn – đặc trưng cho mô men trượt cạnh. Mỗi một hệ số đặc trưng cho một chiều tự do trong chuyển động của vật thể. Thông thường, dải hoạt động của vật thể là lớn, do đó, kích thước của cơ sở dữ liệu cũng lớn, nên số lượng trường hợp cần tính toán là rất nhiều. Với phương pháp truyền thống, việc sinh cơ sở dữ liệu hệ số tốn nhiều thời gian, công sức và thường xuất hiện lỗi chủ quan do người thực hiện. Phương pháp và hệ thống mới sinh ra để giải quyết các vấn đề này.

Tham chiếu đến Hình 3, phương pháp tự động tìm kiếm, phân tích và tổng hợp dữ liệu được thực hiện qua các bước sau: bước 1: thu thập dữ liệu đầu ra, cụ thể dữ liệu đầu ra từ chương trình tính toán động lực học chất lỏng; bước 2: phân loại và tìm kiếm tệp dữ

liệu hệ số; bước 3: lọc và xử lý dữ liệu hệ số; bước 4: chuẩn hóa dữ liệu hệ số; bước 5: sinh cơ sở dữ liệu hệ số động lực học. Cụ thể như sau:

Bước 1: thu thập dữ liệu đầu ra chương trình tính toán động lực học chất lỏng;

Đầu ra của một trường hợp tính toán động lực học chất lỏng là một tệp chứa tất cả các thông tin quan trọng như hình học, lưới rời rạc, thiết lập tính toán, kết quả tính toán... Để xây dựng cơ sở dữ liệu hệ số, cần sử dụng rất nhiều trường hợp tính toán. Do đó, số lượng tệp đầu ra từ chương trình tính toán là cực kỳ lớn và thường được lưu trữ phân tán, không có thứ tự. Vì vậy, mục đích của bước đầu tiên là thu thập tất cả các tệp thông tin và sắp xếp vào một tệp duy nhất. Như vậy, đầu vào của bước này là các tệp thông tin của các trường hợp tính toán bị lưu trữ phân tán, đầu ra là một tệp tin duy nhất chứa tất cả các tệp thông tin tính toán được sắp xếp theo thứ tự thời gian phục vụ quá trình xử lý sau này.

Tại bước này, thực hiện thu thập tất cả các tệp đầu ra từ chương trình tính toán động lực học chất lỏng của tất cả các trường hợp tính toán thành một tệp tổng hợp duy nhất. Việc tổng hợp có thể được thực hiện trực tiếp là sao chép (di chuyển) tệp đầu ra của mỗi một trường hợp tính toán đến tệp tổng hợp hoặc gián tiếp đơn giản là ghi nhớ địa chỉ lưu trữ tệp đầu ra. Hiện tại, hệ thống đề xuất sử dụng cách thu thập gián tiếp, ghi nhớ địa chỉ tệp đầu ra. Ngoài ra, các trường hợp được thu thập vào tệp tin tổng hợp được lưu trữ theo thứ tự thời gian tính toán để dễ dàng trong quá trình duyệt tệp tin trong quá trình xử lý và sắp xếp vào cơ sở dữ liệu hệ số sau này (tham chiếu Hình 2). Như vậy, mục đích chính của bước này là thu thập và sắp xếp toàn bộ tệp đầu ra từ chương trình tính toán của tất cả các trường hợp tính toán thành một tệp tổng hợp. Bước thu thập là bước tiền đề, chuẩn bị đầu vào cho các bước xử lý tiếp theo.

Bước 2: phân loại, tìm kiếm tệp dữ liệu hệ số;

Tham chiếu Hình 2, tất cả đầu ra của chương trình tính toán cho các trường hợp tính toán từ số 1 đến số MN (tương ứng với cơ sở dữ liệu hệ số dạng bảng hai chiều bao gồm N hàng, M cột) được thu thập trong một tệp tổng hợp duy nhất. Trong đó, mỗi một tệp đầu ra của chương trình tính toán động lực học chất lỏng cho một trường hợp bao

gồm các loại dữ liệu chính sau: dữ liệu hình học, dữ liệu lưới rời rạc, dữ liệu thiết lập trường hợp tính toán, tham số hệ thống và dữ liệu hệ số. Mỗi một loại dữ liệu thường được lưu trữ tại một tệp có tên đặt bởi chương trình tính toán theo quy ước có sẵn. Việc quy ước đặt tên sẵn tệp được sinh bởi chương trình tính toán có tác dụng phân loại các thông tin trong tệp đầu ra. Ví dụ: trong tệp dữ liệu của trường hợp thứ nhất, sẽ có các tệp con chứa thông tin về hình học, lưới rời rạc...

Đồng thời, dựa theo quy ước đặt tên này, tệp được quan tâm trong bước này là tệp chứa các hệ số đầu ra, được nhận dạng thông qua tên của tệp. Cụ thể, các tệp con chứa các loại dữ liệu tương ứng được kiểm tra bằng thuật toán thử truy cập.

Ví dụ: quy ước đặt tên của tệp chứa hình học là “geo”, để kiểm tra tệp con này có tồn tại hay không trong một tệp đầu ra tính toán thì tiến hành truy cập thử vào tệp “<địa chỉ tệp đầu ra>/geo”. Khi đó, một trường hợp tính toán hợp lệ khi và chỉ khi tệp dữ liệu đầu ra có đầy đủ các tệp con chứa dữ liệu về hình học, lưới rời rạc, thiết lập tính toán, tham số hệ thống và hệ số đầu ra. Khi tệp dữ liệu đầu ra không hợp lệ, thông báo sẽ đưa ra chính xác tên của tệp bị thiếu. Sử dụng thuật toán trên hoàn toàn có thể xác định được tệp con chứa dữ liệu hệ số đầu ra.

Như vậy, đầu vào của bước 2 là tệp dữ liệu tổng hợp các trường hợp tính toán từ bước 1. Sau đó, cấu trúc bên trong của từng trường hợp tính toán được tự động kiểm tra và phân loại bằng thuật toán thử truy cập. Đầu ra sau bước phân loại là tệp con chứa các hệ số động lực học chất lỏng.

Sau khi xác định được tệp chứa các hệ số, sử dụng thuật toán thử truy cập để xác định sự tồn tại các tệp con chứa từng loại hệ số.

Thông thường, chương trình tính toán sinh sáu tệp chứa hệ số bao gồm: Cx, hệ số lực dọc thân; Cy, hệ số lực trượt cạnh; Cz, hệ số lực pháp tuyến; Cl, hệ số mô men quanh trục dọc; Cm, hệ số mô men chúc góc; Cn, hệ số mô men trượt cạnh. Ví dụ: “<địa chỉ tệp đầu ra>/<tệp chứa hệ số>/cx”.

Ngoài ra, trong bước này, phương pháp đề xuất thêm thuật toán kiểm tra về số lượng và chủng loại của các tệp chứa hệ số. Nếu đầu ra của chương trình tính toán không đủ số lượng là sáu và chủng loại (Cx, Cy, Cz, Cl, Cm, Cn) thì sẽ đưa ra thông báo về loại dữ liệu còn thiếu tới người sử dụng và sau đó bỏ qua trường hợp hiện tại để kiểm tra đầu ra của chương trình tính toán cho trường hợp tiếp theo. Qua đó, xác định được số lượng và chủng loại các tệp con chứa trong tệp chứa hệ số. Hơn nữa, tệp tin chứa đúng loại hệ số theo yêu cầu để tổng hợp thành cơ sở dữ liệu được xác định.

Toàn bộ quy trình và thuật toán trên được thực hiện trong mô đun tìm kiếm dữ liệu của hệ thống tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn.

Bước 3: lọc và xử lý dữ liệu;

Tệp chứa dữ liệu của một loại hệ số thường là một danh sách các số sinh bởi chương trình tính toán theo thời gian. Các số này thường biểu diễn một dao động dạng hình sin. Tệp chứa dữ liệu hệ số - đầu ra từ bước 2 được kiểm tra về nội dung để lọc ra các tệp chứa dữ liệu hội tụ, nói cách khác, được kiểm tra về tính hội tụ sử dụng thuật toán hội tụ trong mô đun xử lý dữ liệu của hệ thống đề xuất.

Sau đó, dữ liệu hội tụ này sẽ được xử lý toán học để xác định giá trị đại diện - đầu ra của bước 3. Trước tiên, tệp chứa hệ số được mở và kiểm tra số lượng phần tử bên trong bằng thuật toán duyệt dữ liệu thực hiện từ phần tử đầu tiên cho tới phần tử cuối cùng. Tổng số lượng phần tử cần tương ứng với tham số trong phần thiết lập trường hợp tính toán. Nếu không, thông báo lỗi về ghi dữ liệu của chương trình tính toán chất lỏng sẽ được đưa ra.

Tham chiếu Hình 1 mục “Kiểm tra hội tụ”, khi biểu diễn danh sách các phần tử trong tệp chứa hệ số dưới dạng đồ thị phụ thuộc theo thời gian, kết quả nhận được thường là một hình dao động dạng sin. Dữ liệu được xác định là hội tụ khi và chỉ khi thỏa mãn các điều kiện sau đây: điều kiện thứ nhất biên độ dao động của năm (05) chu kỳ cuối cùng giảm nhiều hơn 10%; điều kiện thứ hai là biên độ dao động của năm (05) chu kỳ cuối không quá 10% giá trị cực tiểu. Các tham số về năm (05) chu kỳ hoặc 10% có thể

được điều chỉnh trong từng tình huống cụ thể. Ngược lại, nếu dữ liệu không chuẩn, phương pháp và hệ thống đề xuất sẽ có thông báo được đưa ra với hai cấp độ như sau: cấp độ thứ nhất là dữ liệu không hội tụ (không thỏa mãn điều kiện thứ nhất và thứ hai); cấp độ thứ hai là dữ liệu cần tính toán thêm (thỏa mãn một trong hai điều kiện). Thông báo này nhằm đưa ra gợi ý cho người sử dụng thay đổi thiết lập trong chương trình tính toán động lực học chất lỏng để thu được dữ liệu hội tụ. Sau bước lọc, các tệp chứa dữ liệu hệ số hội tụ tiếp tục được xử lý bằng toán học để xác định giá trị đại diện theo công thức sau:

$$V_{\text{pt}} = \frac{\sum_{5T} v}{\text{Num}(5T)}$$

Trong đó, giá trị đại diện (V_{pt}) được tính bằng trung bình cộng các giá trị hệ số (v) trong khoảng năm (05) chu kỳ cuối cùng (5T) của tệp dữ liệu hệ số. Giá trị hệ số đại diện này được lấy là xấp xỉ cho toàn bộ tệp chứa dữ liệu hệ số. Sau này, giá trị hệ số đại diện là giá trị sẽ được điền vào cơ sở dữ liệu hệ số động lực học chất lỏng.

Việc đếm số chu kỳ và tính biên độ dao động của dữ liệu hệ số được thực hiện nhờ thuật toán nhận dạng đồ thị. Tham chiếu Hình 4, các giá trị hệ số trong tệp chứa được minh họa bằng các chấm tròn. Khi đó, lượng biến thiên dy được xác định bằng hiệu của hai giá trị liền kề. Khi lượng biến thiên đổi dấu từ dương (+) sang âm (-) thì xác định một cực đại. Ngược lại, khi lượng biến thiên đổi dấu từ âm (-) sang dương (+) thì xác định một cực tiểu. Hiệu của một cực đại và một cực tiểu liền kề xác định biên độ dao động. Thuật toán nhận dạng đồ thị, duyệt dữ liệu và xử lý toán học dữ liệu được thực hiện tại môđun xử lý dữ liệu trong hệ thống đề xuất.

Bước 4: chuẩn hóa dữ liệu;

Trong bước này, giá trị đại diện tính toán từ bước 3 được kiểm tra các điều kiện về giới hạn cực đại, cực tiểu và liên hệ với các giá trị đại diện lân cận trong cơ sở dữ liệu hệ số. Thông thường, mỗi loại dữ liệu hệ số động lực học chất lỏng trên một loại đối tượng vật thể thường có xu thế nhất định (ví dụ: hệ số lực pháp tuyến Cz trên máy bay thường

có một giai đoạn dạng tuyến tính). Dựa vào đó, phương pháp và hệ thống đề xuất có thể kiểm tra và đưa ra thông báo về các điểm không tuân theo quy luật của các điểm lân cận hoặc vượt quá các giới hạn. Bên cạnh đó, giá trị đại diện tính toán trong bước 3 cũng được chuẩn hóa về định dạng (số chữ số sau dấu phẩy, kiểu lưu dữ liệu...) để phù hợp với cơ sở dữ liệu hệ số động lực học chất lỏng.

Cụ thể, giá trị hệ số được so sánh với các giá trị cực trị như sau:

$$\text{Min}(C_i) \leq C_{rpt} \leq \text{Max}(C_i)$$

Trong đó, C_{rpt} là giá trị hệ số đại diện được kiểm tra, i có thể là x, y, z, l, m, n theo yêu cầu từ người dùng. $\text{Min}(C_i)$, $\text{Max}(C_i)$ tương ứng là giá trị cực tiểu và cực đại của từng loại hệ số động lực học chất lỏng. Các giá trị cực trị này là thường được xác định trước thông qua thực nghiệm hoặc kinh nghiệm.

Khi giá trị đại diện nằm trong giới hạn cho phép, xu thế thay đổi của giá trị đó sẽ được kiểm tra bằng cách so sánh với các giá trị lân cận. Xu thế thay đổi được xác định từ trước và cho từng loại hệ số thông qua các kết quả nghiên cứu lý thuyết và kinh nghiệm. Ví dụ: xu thế thay đổi của hệ số lực pháp tuyến $C_z(x)$ là tăng dần trong khoảng $x = [0 \div 10]$, khi đó, giá trị hệ số phải thỏa mãn điều kiện sau:

$$C_z(x=4) \leq C_{rpt}(x=5) \leq C_z(x=6)$$

Khi đó, giá trị đại diện C_{rpt} tương ứng với trường hợp $x = 5$ phải nằm giữa hai giá trị lân cận.

Nếu giá trị hệ số đại diện không thỏa mãn điều kiện giới hạn và xu thế thay đổi, thông báo giá trị đại diện vượt giới hạn hoặc không thỏa mãn xu thế thay đổi được đưa ra. Sau đó, người dùng có hai lựa chọn để tiếp tục chương trình: lựa chọn thứ nhất là loại bỏ giá trị đại diện – giá trị đại diện bị hủy, bỏ qua trường hợp hiện tại chuyển qua trường hợp kế tiếp; lựa chọn thứ hai là chấp nhận giá trị đại diện.

Trong hệ thống đề xuất, định dạng của cơ sở dữ liệu hệ số là 04 chữ số thập phân sau dấu phẩy. Bước này chủ yếu sử dụng các biến đổi toán học để chuẩn hóa lại giá trị hệ

số đại diện. Cụ thể, việc chuẩn hóa số lượng chữ số sau dấu phẩy được thực hiện bằng thuật toán làm tròn đơn giản như sau:

$$C_{rpt} = \text{int}(1000 * C_{rpt}) / 1000$$

Như vậy, dữ liệu đầu ra của bước chuẩn hóa dữ liệu là hệ số đại diện tính trong bước 3 đã được chuẩn về định dạng (04 chữ số thập phân sau dấu phẩy) và chuẩn về giới hạn và xu hướng thay đổi. Bước chuẩn hóa dữ liệu là phần đầu tiên trong môđun phân phối dữ liệu của hệ thống đề xuất.

Bước 5: sinh cơ sở dữ liệu hệ số động lực học chất lỏng;

Cơ sở dữ liệu động lực học chất lỏng là thành phần quan trọng trong việc xây dựng mô hình động lực học các vật thể hoạt động trong môi trường không khí, nước hoặc không khí và nước. Cơ sở dữ liệu này đặc trưng cho đáp ứng của vật thể trong các trạng thái hoạt động khác nhau. Vì vậy, nó là thành phần tiên quyết đến độ chân thực của các hệ thống mô phỏng.

Cụ thể, cấu trúc tổ chức của cơ sở dữ liệu hệ số động lực học chất lỏng thường sử dụng các bảng hai chiều hoặc ba chiều để lưu trữ sáu hệ số động lực học. Giá trị đại diện sau khi được tính toán từ bước 3, chuẩn hóa trong bước 4 sẽ được lưu trữ tại cơ sở dữ liệu này tại vị trí xác định. Vị trí lưu trữ được xác định thông qua thứ tự của nó trong tệp tổng hợp dữ liệu trong bước 1.

Tham chiếu Hình 7, có ba trường hợp xảy ra.

Trường hợp thứ nhất, cơ sở dữ liệu chứa hệ số động lực học chưa được khởi tạo. Trong trường hợp này, cơ sở dữ liệu được khởi tạo theo định dạng chuẩn và giá trị đại diện chuẩn được lưu trữ tại vị trí đã xác định.

Trường hợp thứ hai, cơ sở dữ liệu có nhưng sai về định dạng (số hàng, số cột, số chiều...). Khi đó, hệ thống và phương pháp đề xuất sẽ đưa ra thông báo về lỗi sai để người sử dụng sửa chữa, khắc phục.

Trường hợp thứ ba, cơ sở dữ liệu đã tồn tại và đúng chuẩn. Trong trường hợp này, vị trí mà giá trị đại diện dự định được lưu trữ sẽ được kiểm tra. Nếu đã tồn tại giá trị thì đưa ra thông báo và cho phép thay đổi hoặc giữ nguyên. Trong trường hợp vị trí lưu trữ chưa tồn tại giá trị nào thì giá trị đại diện chuẩn sẽ được lưu tại đây.

Như vậy, đầu vào của bước 5 là các hệ số đại diện đã được chuẩn hóa theo định dạng của cơ sở dữ liệu hệ số, đầu ra là cơ sở dữ liệu của loại hệ số yêu cầu. Trong hệ thống đề xuất, môđun phân phối dữ liệu là phần thực hiện sinh cơ sở dữ liệu hệ số dựa trên các chỉ dẫn trong bước 5. Việc kiểm tra cơ sở dữ liệu cũng được thực hiện bằng thuật toán thử truy cập như trong bước 1.

Tham chiếu Hình 3, để thực hiện phương pháp đề xuất, cụ thể là phương pháp và hệ thống tự động tìm kiếm, phân tích và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng cần thực hiện năm bước. Đồng thời, phương pháp đề xuất cần sử dụng hệ thống bao gồm ba môđun. Đầu tiên, dữ liệu đầu ra của tất cả các trường hợp tính toán được môđun tìm kiếm dữ liệu phân loại và tìm kiếm tệp chứa dữ liệu hệ số theo yêu cầu từ người dùng. Tiếp theo, tệp chứa dữ liệu được kiểm tra tiêu chuẩn hội tụ và xác định giá trị đại diện như trình bày trong bước 3. Cuối cùng, môđun phân phối dữ liệu tiến hành chuẩn hóa dữ liệu và lưu trữ vào đúng vị trí trong cơ sở dữ liệu hệ số. Cụ thể, các môđun được thực thi như sau:

Tham chiếu Hình 5, môđun tìm kiếm dữ liệu có mục đích thực hiện bước 2 để phân loại và tìm kiếm tệp chứa hệ số. Cụ thể, hệ thống đề xuất sẽ duyệt các tệp đầu ra từ chương trình tính toán gồm tất cả các trường hợp được thu thập tại một tệp dữ liệu tổng hợp. Bước kiểm tra thứ nhất, kiểm tra về cấu trúc của các tệp sinh bởi chương trình tính toán nhằm đảm bảo tất cả các trường hợp tính toán đều sinh ra hệ số khí động và đủ số lượng hệ số khí động là 06. Bên cạnh đó, bước kiểm tra thứ hai, kiểm tra về chủng loại là kiểm tra về thành phần của dữ liệu hệ số động lực học chất lỏng sinh ra bởi chương trình tính toán. Sau khi cấu trúc dữ liệu đầu ra là chuẩn xác, tệp chứa dữ liệu hệ số được tìm kiếm theo chủng loại yêu cầu từ người dùng. Ví dụ: người dùng muốn tổng hợp cơ sở dữ

liệu hệ số C_x , môđun sẽ thực hiện tìm kiếm các tệp chứa hệ số tương ứng theo tên được đặt theo quy ước như “ cx ”.

Tham chiếu Hình 5, môđun tìm kiếm dữ liệu có tích hợp thuật toán thử truy cập và đưa ra thông báo khi phát hiện lỗi: thứ nhất là lỗi về cấu trúc bên trong của tệp tin đầu ra từ chương trình tính toán; thứ hai là lỗi về số lượng, chủng loại của tệp con chứa dữ liệu hệ số. Tham chiếu Hình 8 – minh họa cấu trúc cơ bản của tệp đầu ra chương trình tính toán chất lỏng, thuật toán truy cập thử dựa trên tên và cách sắp xếp các loại tệp dữ liệu được sinh bởi chương trình tính toán.

Ví dụ: môđun tìm kiếm dữ liệu sử dụng thuật toán thử truy cập để lần lượt kiểm tra các đường dẫn sau:

(1) *.../tinhchatlong(i)/hinhhoc*

(2) *.../tinhchatlong(i)/luoi*

(3) *.../tinhchatlong(i)/thietlap*

(4) *.../tinhchatlong(i)/hethong*

(5) *.../tinhchatlong(i)/heso*

Trong đó, (i) là số của trường hợp tính toán. Nếu đường dẫn “*.../tinhchatlong(i)/hinhhoc*” tồn tại, thuật toán trả ra giá trị 1, chứng tỏ tệp dữ liệu đầu ra có tệp con mang tên “*hinhhoc*”. Ngược lại, nếu đường dẫn không tồn tại, thuật toán trả ra giá trị 0, chứng tỏ tệp dữ liệu có tên “*hinhhoc*” không có trong tệp “*tinhchatlong(i)*”. Khi đó, hệ thống đưa ra thông báo không tồn tại tệp có tên “*hinhhoc*” và dùng khảo sát trường hợp tính toán (i) để chuyển sang kiểm tra trường hợp tính toán $(i+1)$. Thuật toán thử truy cập được sử dụng để kiểm tra sự tồn tại của các tệp khác theo cách tương tự. Đầu ra của môđun tìm kiếm dữ liệu là một trong sáu tệp chứa dữ liệu hệ số sau:

(1) *.../tinhchatlong(i)/heso/cx*

(2) *.../tinhchatlong(i)/heso/cy*

(3) *.../tinhchatlong(i)/heso/cz*

(4) *.../tinhchatlong(i)/heso/cl*

(5).../tinhchatlong(i)/heso/cm

(6).../tinhchatlong(i)/heso/cn

Tham chiếu Hình 6, môđun xử lý dữ liệu được xây dựng để thực hiện hai mục đích trong bước 3 của phương pháp đề xuất. Trong đó, mục đích thứ nhất là kiểm tra về số lượng dữ liệu hệ số động lực học chứa trong tệp chứa hệ số. Số lượng dữ liệu này phải tương ứng với tham số của chương trình tính toán. Ví dụ: thiết lập chương trình tính toán là 2000 vòng lặp thì số lượng phần tử trong tệp chứa hệ số tương ứng là 2000. Cụ thể, quá trình thực hiện của chương trình tính toán chất lỏng bao gồm nhiều vòng lặp, kết thúc mỗi vòng lặp thì kết quả hệ số được lưu tại một dòng trong tệp tin chứa hệ số. Khi đó, kiểm tra số dòng của tệp tin chứa hệ số sẽ xác định được số lượng dữ liệu.

Tệp chứa dữ liệu hệ số là một tập hợp các giá trị sinh ra trong quá trình tính toán. Khi biểu diễn dưới dạng đồ thị có trục hoành là thời gian và trục tung là giá trị hệ số thì đồ thị thường có dạng dao động hình sin (tham chiếu Hình 4). Tính hội tụ của tệp chứa hệ số được kiểm tra theo các tiêu chuẩn trình bày trong bước 3. Thuật toán sử dụng để kiểm tra tính hội tụ của đồ thị là thuật toán nhận dạng đồ thị. Cụ thể, tham chiếu Hình 4, lượng biến thiên dy được xác định bằng hiệu của hai giá trị liên kề. Khi lượng biến thiên đổi dấu từ dương (+) sang âm (-) thì xác định một cực đại. Ngược lại, khi lượng biến thiên đổi dấu từ âm (-) sang dương (+) thì xác định một cực tiểu. Hiệu của một cực đại và một cực tiểu liên kề xác định biên độ dao động. Lượng biến thiên dy đổi dấu bốn lần thì xác định một chu kỳ. Căn cứ vào tiêu chuẩn hội tụ: biên độ dao động của năm (05) chu kỳ cuối cùng giảm nhiều hơn 10%; biên độ dao động của năm (05) chu kỳ cuối không quá 10% giá trị cực tiểu.

$$\frac{A_{N-5} - A_N}{A_{N-5}} \geq 10\%$$

$$\frac{A_{N-5} + A_{N-4} + A_{N-3} + A_{N-2} + A_{N-1} + A_N}{5} \leq 10\% \times \text{Min}(T-5)$$

Trong đó, A_i là biên độ tại chu kỳ thứ i , N là chu kỳ cuối cùng dữ liệu và $\text{Min}(T-5)$ là giá trị cực tiểu tại chu kỳ thứ $(T-5)$.

Khi dữ liệu đã hội tụ thì sẽ được môđun xử lý dữ liệu tính giá trị đại diện bằng công thức toán học trình bày trong bước 3. Giá trị này sẽ được chuẩn hóa và lưu trữ trong cơ sở dữ liệu hệ số.

Ngược lại, môđun sẽ đưa ra thông báo về các lỗi sai như: thiếu số lượng hệ số, dữ liệu không hội tụ. Sau đó, môđun sẽ bỏ qua tệp dữ liệu hệ số này để chuyển qua khảo sát trường hợp tính toán khác.

Tổng hợp, môđun thực hiện các chức năng: kiểm tra tệp chứa hệ số, kiểm tra tính hội tụ, tính toán giá trị đại diện và đưa ra các thông báo theo trình tự trong Hình 6. Mỗi chức năng trên được tổ chức thành một hàm riêng biệt thực hiện các phép tính toán học và thuật toán trình bày trên nhằm mục đích thực hiện bước 3 trong phương pháp tự động tìm kiếm, phân tích và tổng hợp dữ liệu lớn.

Tham chiếu hình 7, môđun phân phối dữ liệu được xây dựng để chuẩn hóa định dạng giá trị được tính toán trong môđun xử lý dữ liệu cho phù hợp với định dạng trong cơ sở dữ liệu hệ số động lực học chất lỏng (tương ứng với bước 4 trong phương pháp đề xuất). Sau đó, vị trí của giá trị hệ số chuẩn hóa trong cơ sở dữ liệu được xác định dựa vào thứ tự của nó trong tệp thu thập dữ liệu tại bước 1. Ví dụ: bảng hệ số là bảng hai chiều thì vị trí hàng và cột của dữ liệu được tìm kiếm như sau:

$$Column = \text{int}(i / M) + 1$$

$$Row = \text{mod}(i / M) + 1$$

Trong đó, i là thứ tự của dữ liệu trong tệp tổng hợp, M là số hàng của bảng hệ số dữ liệu. int là phép toán lấy phần nguyên của phép chia i cho M . mod là phép toán lấy phần dư của phép chia i cho M .

Môđun này cũng cho phép kiểm trùng dữ liệu khi đã tồn tại giá trị ở vị trí lưu hoặc khởi tạo cơ sở dữ liệu khi nó chưa tồn tại. Cụ thể, môđun dùng thuật toán thử truy cập để kiểm tra sự tồn tại của cơ sở dữ liệu hệ số. Nếu cơ sở dữ liệu không tồn tại, chúng tôi trường hợp đang xử lý là trường hợp đầu tiên, môđun sẽ đưa ra thông báo cho phép người dùng tạo một tệp dữ liệu mới để lưu trữ các hệ số. Khi cơ sở dữ liệu đã tồn tại, môđun

này cho phép truy cập vào bảng hệ số khí động tại vị trí xác định từ phần trên để kiểm tra giá trị. Nếu giá trị đọc được khác “*Null*”, chứng tỏ đã có giá trị lưu trữ tại vị trí trên. Khi đó, môđun sẽ đưa ra thông báo cho phép người dùng thay thế hoặc giữ nguyên giá trị cũ.

Cuối cùng, dữ liệu chuẩn hóa được lưu tại vị trí đã được xác định trong cơ sở dữ liệu như trình bày trên. Như vậy, môđun phân phối dữ liệu có cấu tạo gồm bốn hàm con thực hiện các chức năng tương ứng bao gồm: hàm thứ nhất là hàm chuẩn hóa dữ liệu; hàm thứ hai là hàm kiểm tra cơ sở dữ liệu hệ số; hàm thứ ba là hàm xác định vị trí của hệ số trong cơ sở dữ liệu; hàm thứ tư là hàm kiểm trùng và lưu vào cơ sở dữ liệu nhằm thực hiện các bước 4 và 5 của phương pháp đề xuất.

Yêu cầu bảo hộ

1. Phương pháp tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng gồm năm bước:

bước 1: thu thập dữ liệu đầu ra chương trình tính toán động lực học chất lỏng trong đó thực hiện thu thập tất cả các tệp đầu ra từ chương trình tính toán động lực học chất lỏng của tất cả các trường hợp tính toán thành một tệp tổng hợp duy nhất bằng phương pháp gián tiếp, ghi nhớ địa chỉ tệp đầu ra;

bước 2: phân loại và tìm kiếm dữ liệu, cụ thể các tệp dữ liệu đầu ra từ chương trình tính toán động lực học chất lỏng được đảm bảo về số lượng và chủng loại;

một tệp đầu ra của chương trình tính toán động lực học chất lỏng có tên đặt bởi chương trình tính toán theo quy ước có sẵn có tác dụng phân loại các thông tin trong tệp đầu ra;

các tệp con chứa các loại dữ liệu tương ứng được kiểm tra bằng thuật toán thử truy cập bằng cách tiến hành các trường hợp tính toán truy cập thử vào tệp "<địa chỉ tệp đầu ra>/<tên quy ước>"; thu được tệp con chứa các hệ số động lực học chất lỏng;

sử dụng thuật toán thử truy cập xác định các tệp con chứa từng loại hệ số ứng với 06 tệp chứa hệ số bao gồm: Cx, hệ số lực dọc thân; Cy, hệ số lực trượt cạnh; Cz, hệ số lực pháp tuyến; Cl, hệ số mô men quanh trục dọc; Cm, hệ số mô men chúc góc; Cn, hệ số mô men trượt cạnh;

bước 3: lọc và xử lý dữ liệu hệ số, cụ thể tệp chứa dữ liệu hệ số - đầu ra từ bước 2 được kiểm tra về nội dung để lọc ra các tệp chứa dữ liệu hội tụ bằng cách biểu diễn danh sách các phần tử trong tệp chứa hệ số dưới dạng đồ thị phụ thuộc theo thời gian và xác định mức độ thỏa mãn các yếu tố biên độ dao động của năm (05) chu kỳ cuối cùng giảm nhiều hơn 10% và biên độ dao động của năm (05) chu kỳ cuối không quá 10% giá trị cực tiểu;

dữ liệu hội tụ này sẽ được xử lý toán học để xác định giá trị đại diện;

bước 4: chuẩn hóa dữ liệu, trong đó giá trị đại diện được chuẩn hóa định dạng theo các tiêu chuẩn về cận trên, cận dưới và liên hệ với các giá trị lân cận trong cơ sở dữ liệu hệ số động lực học chất lỏng;

bước 5: sinh cơ sở dữ liệu hệ số động lực học chất lỏng, trong đó cơ sở dữ liệu loại hệ số được yêu cầu được xây dựng dưới dạng các bảng hai chiều hoặc ba chiều từ các giá trị hệ số đại diện của các trường hợp tính toán khác nhau; loại bảng và vị trí lưu các giá trị đại diện phụ thuộc vào loại hệ số và trường hợp tính toán.

2. Phương pháp tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng theo điểm 1, trong đó tại bước 2, thuật toán thử truy cập hợp lệ khi và chỉ khi tệp dữ liệu đầu ra có đầy đủ các tệp con chứa dữ liệu về hình học, lưới rời rạc, thiết lập tính toán, tham số hệ thống và hệ số đầu ra; khi tệp dữ liệu đầu ra không hợp lệ, thông báo sẽ đưa ra chính xác tên của tệp bị thiếu.

3. Phương pháp tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng theo điểm 1, trong đó tại bước 3, việc xác định dữ liệu hội tụ bằng xác định mức độ thỏa mãn các yếu tố đã nêu, các tham số về năm (05) chu kỳ hoặc 10% có thể được điều chỉnh trong từng tình huống cụ thể.

4. Phương pháp tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng theo điểm 1, trong đó tại bước 3, việc đếm số chu kỳ và tính biên độ dao động của dữ liệu hệ số được thực hiện nhờ thuật toán nhận dạng đồ thị, các giá trị hệ số trong tệp chứa được minh họa bằng các chấm tròn, khi đó, lượng biến thiên được xác định bằng hiệu của hai giá trị liền kề;

khi lượng biến thiên đổi dấu từ dương (+) sang âm (-) thì xác định một cực đại;

ngược lại, khi lượng biến thiên đổi dấu từ âm (-) sang dương (+) thì xác định một cực tiểu;

hiệu của một cực đại và một cực tiểu liền kề xác định biên độ dao động.

5. Phương pháp tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng theo điểm 1, trong đó tại bước 5, có ba trường hợp xảy ra

trường hợp thứ nhất, cơ sở dữ liệu chứa hệ số động lực học chưa được khởi tạo, cơ sở dữ liệu được khởi tạo theo định dạng chuẩn và giá trị đại diện chuẩn được lưu trữ tại vị trí đã xác định;

trường hợp thứ hai, cơ sở dữ liệu có nhưng sai về định dạng (số hàng, số cột, số chiều...), khi đó, hệ thống và phương pháp đề xuất sẽ đưa ra thông báo về lỗi sai để người sử dụng sửa chữa, khắc phục;

trường hợp thứ ba, cơ sở dữ liệu đã tồn tại và đúng chuẩn, vị trí mà giá trị đại diện dự định được lưu trữ sẽ được kiểm tra; nếu đã tồn tại giá trị thì đưa ra thông báo và cho phép thay đổi hoặc giữ nguyên; trong trường hợp vị trí lưu trữ chưa tồn tại giá trị nào thì giá trị đại diện chuẩn sẽ được lưu tại đây.

6. Hệ thống tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng bao gồm ba mô đun sau:

mô đun tìm kiếm dữ liệu có tích hợp thuật toán thử truy cập và đưa ra thông báo khi phát hiện lỗi, thực hiện chức năng tổng hợp, kiểm tra cấu trúc các tệp đầu ra của chương trình tính toán, sau đó tệp dữ liệu đầu ra được phân loại và đầu ra của mô đun tìm kiếm dữ liệu là một trong sáu tệp chứa dữ liệu hệ số; mô đun tìm kiếm dữ liệu đưa ra thông báo khi phát hiện lỗi tương ứng với hai loại lỗi có thể xảy ra: lỗi về cấu trúc tệp tin đầu ra của chương trình tính toán; lỗi về số lượng, chủng loại của tệp con chứa dữ liệu hệ số; cụ thể:

với loại lỗi thứ nhất, cấu trúc đầu ra chuẩn của chương trình tính toán chuẩn bao gồm năm tệp chứa thông tin về: hình dạng vật thể (geo), lưới rời rạc (mesh), thiết lập tính toán (compt), tham số hệ thống khi chạy (log) và dữ liệu hệ số đầu ra (coeff); lỗi cấu trúc tệp là lỗi khi đầu ra của chương trình tính toán không đủ năm loại tệp trên; các tệp này được đặt tên theo quy ước nên thực hiện kiểm tra bằng phương pháp truy cập thử, nếu không tồn tại tên của năm tệp nêu trên thì được thông báo là sai lỗi cấu trúc tệp;

với loại lỗi thứ hai, lỗi này được kiểm tra sau khi kiểm tra lỗi thứ nhất; cụ thể, trong tệp hệ số đầu ra chuẩn (không lỗi) bao gồm sáu tệp con có tên quy

định sẵn (cx, cy, cz, cl, cm, cn); sử dụng phương pháp truy cập thử tương tự như trên để kiểm tra số lượng và chủng loại theo tên của các tệp con; nếu không tồn tại tên của sáu tệp con nêu trên thì được thông báo là sai lỗi về số lượng và chủng loại tệp con chứa hệ số dữ liệu;

môđun xử lý dữ liệu thực hiện các chức năng: kiểm tra tệp chứa hệ số, kiểm tra tính hội tụ, tính toán giá trị đại diện và đưa ra các thông báo; mỗi chức năng trên được tổ chức thành một hàm riêng biệt thực hiện các phép tính toán học và thuật toán nhất định, cụ thể sử dụng thuật toán sử dụng để kiểm tra tính hội tụ của đồ thị;

môđun phân phối dữ liệu có cấu tạo gồm bốn hàm con thực hiện các chức năng tương ứng bao gồm: hàm thứ nhất là hàm chuẩn hóa dữ liệu; hàm thứ hai là hàm kiểm tra cơ sở dữ liệu hệ số; hàm thứ ba là hàm xác định vị trí của hệ số trong cơ sở dữ liệu; hàm thứ tư là hàm kiểm trùng và lưu vào cơ sở dữ liệu, cụ thể, môđun được tích hợp thuật toán thử truy cập để kiểm tra sự tồn tại của cơ sở dữ liệu hệ số; trong đó:

tại hàm chuẩn hóa dữ liệu: dữ liệu trong cơ sở dữ liệu hệ số được quy định chặt chẽ về định dạng như sau: sử dụng bốn bit để lưu trữ giá trị của một hệ số; kiểu dữ liệu trong ngôn ngữ lập trình: float; lưu bốn chữ số thập phân sau dấu phẩy; cụ thể, việc thực hiện ba bước chuẩn hóa dữ liệu này sử dụng các chức năng cơ bản của kỹ thuật lập trình cơ bản;

tại hàm kiểm tra cơ sở dữ liệu hệ số: mục đích nhằm phát hiện, cảnh báo các giá trị bất thường của hệ số trong cơ sở dữ liệu; giá trị được coi là bất thường nếu chúng nằm trong một các trường hợp sau: giá trị hệ số vượt quá giới hạn trên, giá trị hệ số nhỏ hơn giới hạn dưới, giá trị hệ số không tuân theo quy luật; cụ thể, giá trị bất thường về quy luật được hàm kiểm tra cơ sở dữ liệu hệ số kiểm tra thông qua so sánh tương quan với điểm liền kề trái và liền kề phải.

7. Hệ thống tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng theo điểm 6, môđun xử lý dữ liệu sử dụng thuật toán sử dụng để kiểm tra tính hội tụ của đồ thị và tính giá trị đại diện theo công thức sau:

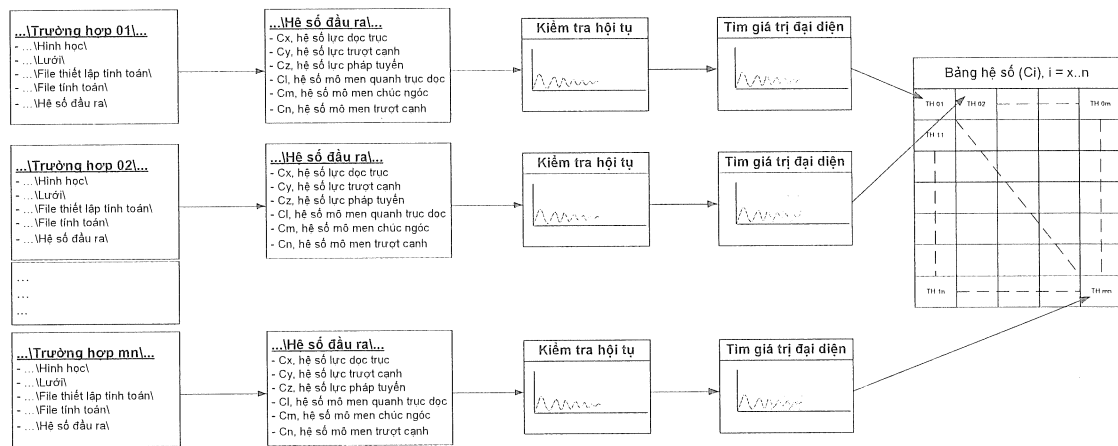
$$V_{rpt} = \frac{\sum_{5T} v}{Num(5T)}$$

trong đó, giá trị đại diện (V_{rpt}) được tính bằng trung bình cộng các giá trị hệ số (v) trong khoảng năm (05) chu kỳ cuối cùng (5T) của tệp dữ liệu hệ số.

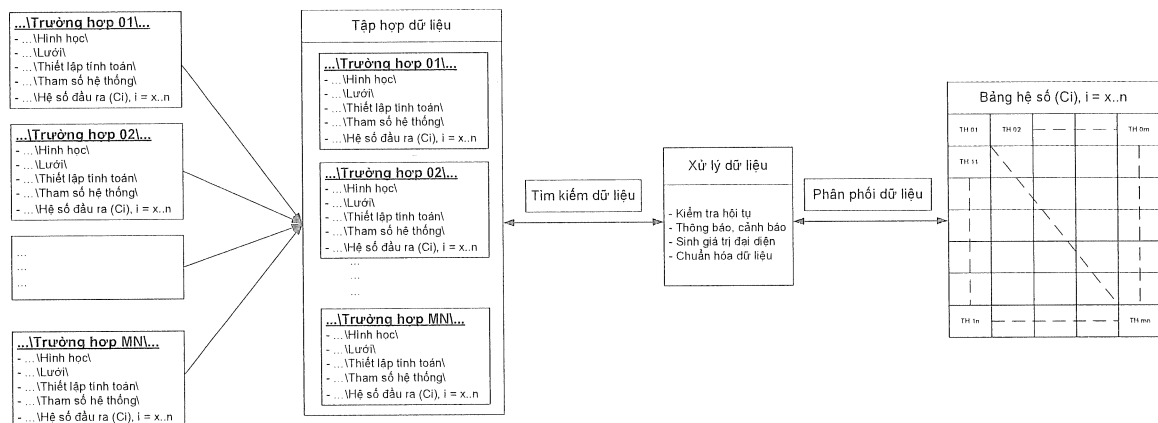
8. Hệ thống tự động tìm kiếm, xử lý và tổng hợp dữ liệu lớn trong bài toán hậu xử lý tính toán động lực học chất lỏng theo điểm 6, trong đó, môđun phân phối dữ liệu

xác định vị trí của giá trị hệ số chuẩn hóa trong cơ sở dữ liệu dựa vào thứ tự của nó trong tệp thu thập dữ liệu,

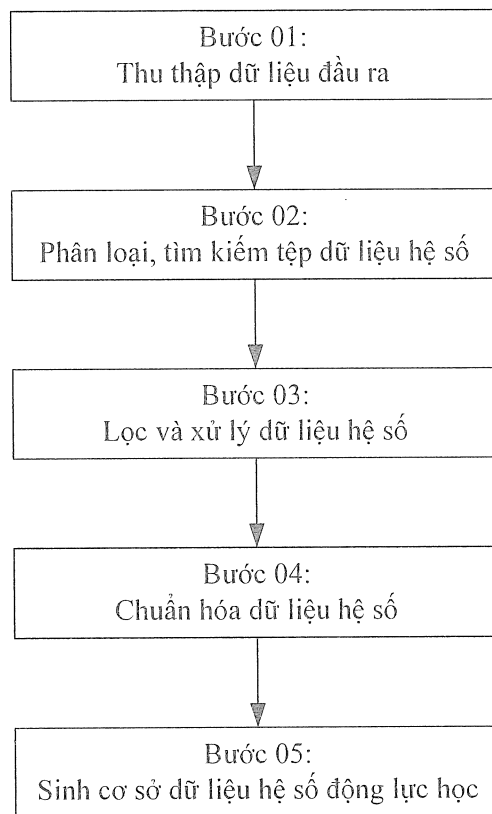
dùng thuật toán thử truy cập để kiểm tra sự tồn tại của cơ sở dữ liệu hệ số.



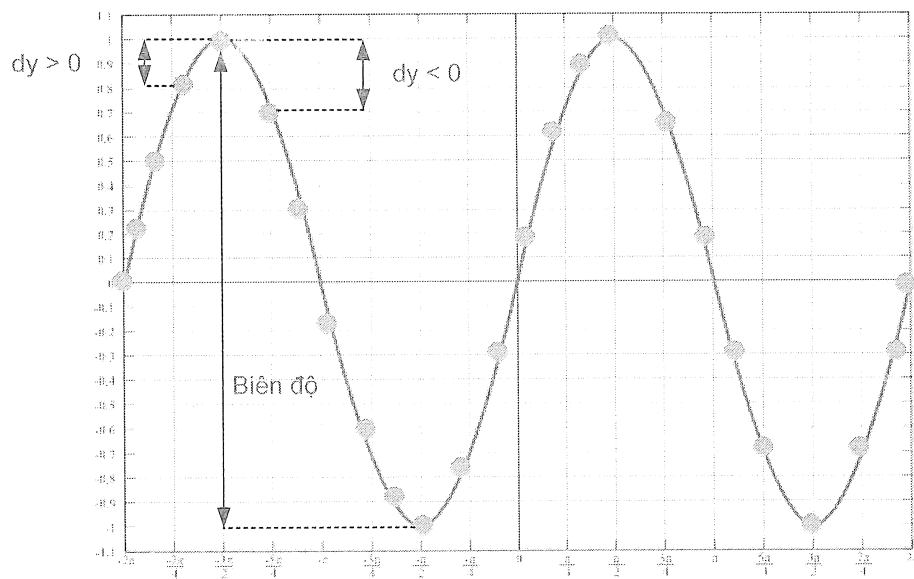
Hình 1



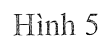
Hình 2

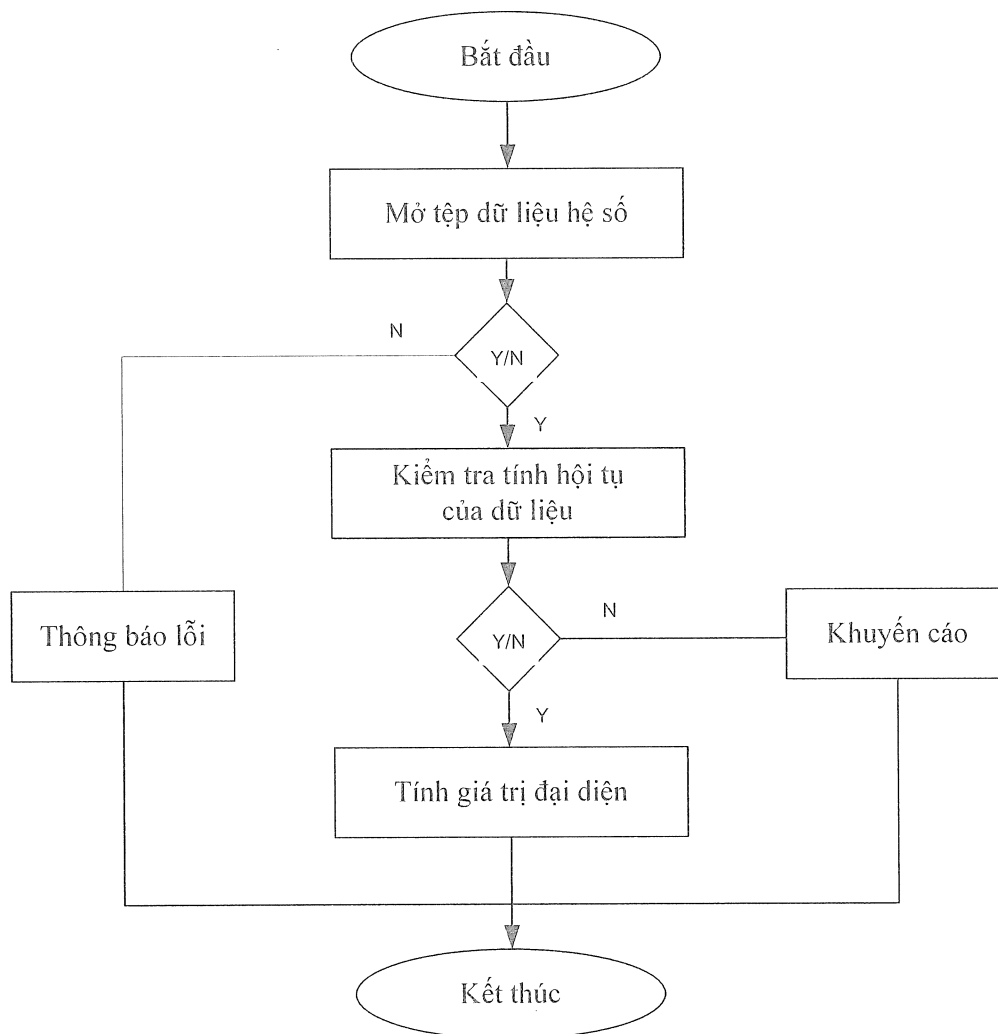


Hình 3

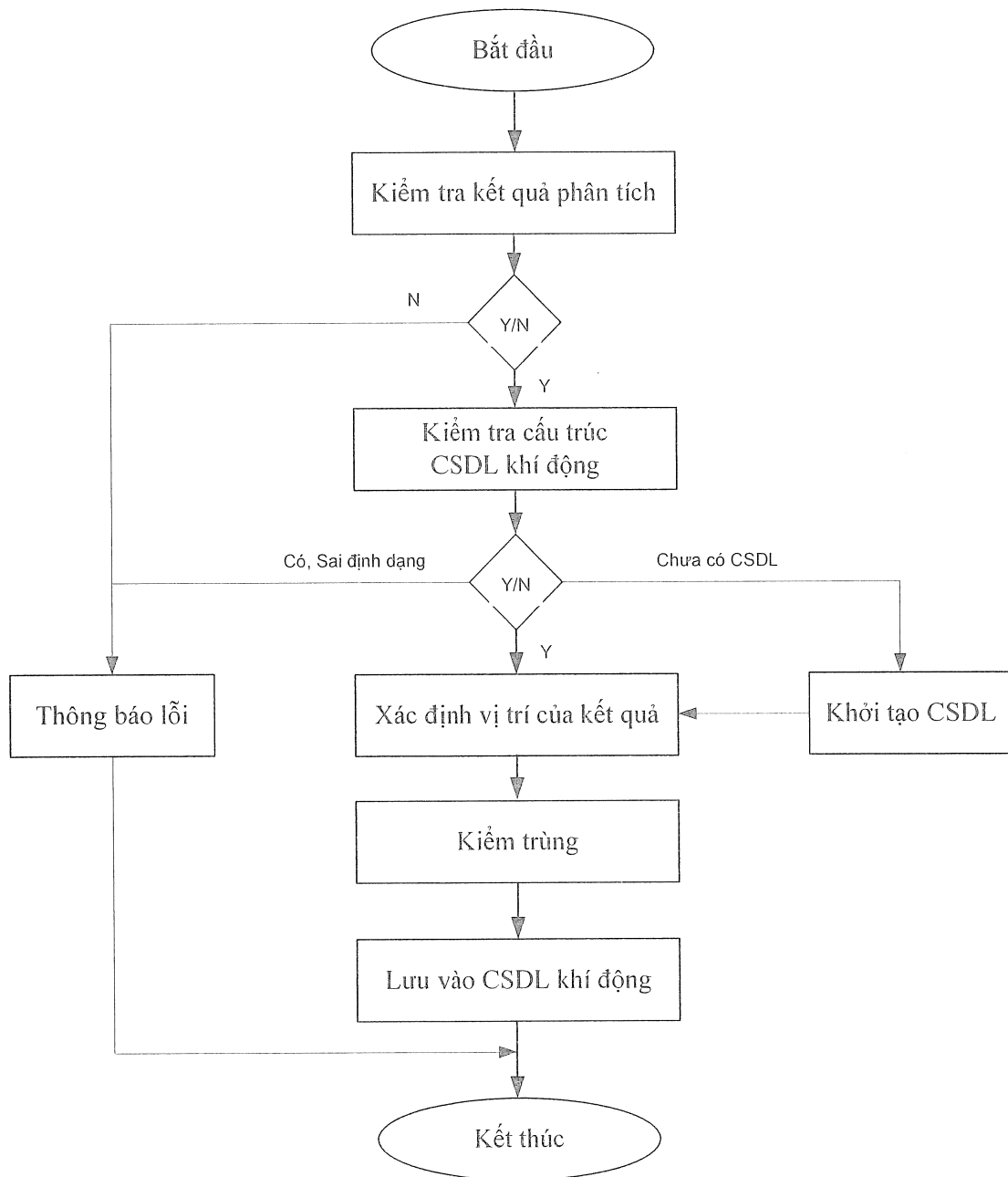


Hình 4

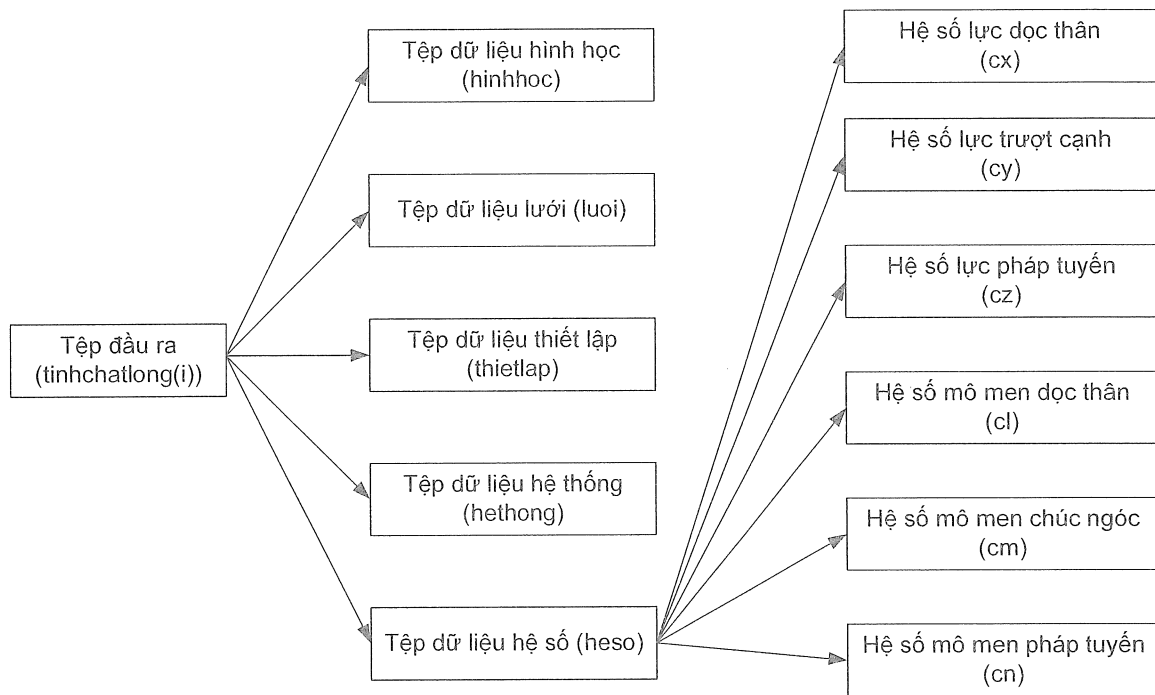




Hình 6



Hình 7



Hình 8