**BÀI BÁO**

**Người dịch: Nguyễn Quang Thái**

*Một hàm niềm tin cơ sở mở rộng trong lý thuyết bằng chứng Dempster-Shafer và nó ứng dụng trong hợp nhất dữ liệu xung đột*

**Tóm lược:** Lý thuyết bằng chứng Dempster–Shafer đã được áp dụng rộng rãi trong lĩnh vực hợp nhất thông tin. Tuy nhiên, khi dữ liệu bằng chứng được thu thập có tính xung đột cao, quy tắc kết hợp Dempster (Dempster combination rule - DCR) không tạo ra kết quả trực quan trong hầu hết thời gian. Để giải quyết vấn đề này, cơ sở hàm niềm tin được đề xuất để sửa đổi phép gán xác suất cơ bản (basic probability assignment - BPA) trong khung đầy đủ nhận thức (frame of discernment - FOD). Tuy nhiên, trong FOD không đầy đủ, giá trị hàm khối lượng của tập rỗng khác không, điều này làm cho hàm niềm tin cơ sở không còn áp dụng được nữa. Trong bài báo này, xem xét các ảnh hưởng của kích thước của FOD và giá trị hàm khối lượng của tập rỗng, một hàm niềm tin mới được đặt tên là hàm niềm tin cơ sở mở rộng (extended base belief function - EBBF) được đề xuất. Phương pháp này có thể sửa đổi BPA trong FOD không đầy đủ và thu được kết quả hợp nhất trực quan bằng cách tính đến các đặc điểm của FOD không đầy đủ. Ngoài ra, EBBF có thể suy biến thành hàm niềm tin cơ sở trong FOD đầy đủ. Đồng thời, bằng cách tính toán entropy niềm tin của BPA đã sửa đổi, chúng tôi thấy rằng giá trị của niềm tin entropy cao hơn trước. Niềm tin entropy được sử dụng để đo lường sự không chắc chắn thông tin, có thể hiển thị xung đột một cách trực quan hơn. Sự gia tăng giá trị của entropy niềm tin là hệ quả của xung đột. This paper also designs an improved conflict data management method based on the EBBF to verify the rationality and effectiveness of the proposed method.

**Từ khóa:**

* Lý thuyết bằng chứng Dempster–Shafer (D-S)
* Chức năng niềm tin
* Khung không đầy đủ của nhận thức (FOD)
* Chức năng niềm tin cơ sở
* Bằng chứng mâu thuẫn
* Quy tắc kết hợp tổng quát

1. **Giới thiệu**

Với sự phát triển của công nghệ máy tính, Internet và các lĩnh vực liên quan khác, công nghệ tổng hợp thông tin, được sinh ra trong lĩnh vực quân sự [1], đã chạy qua mọi ngóc ngách của sản xuất và đời sống của nhân dân [2]. Việc thực hiện hợp nhất thông tin cần phải đối phó với rất nhiều thông tin không chắc chắn. Các công cụ lý thuyết hiện có để đối phó với sự không chắc chắn thông tin bao gồm lý thuyết xác suất [3], lý thuyết tập mờ [4], lý thuyết bằng chứng Dempster–Shafer [5,6], lý thuyết entropy thông tin [7,8], v.v. Lý thuyết bằng chứng Dempster–Shafer được sử dụng rộng rãi và công cụ điển hình để xử lý thông tin không chắc chắn và hợp nhất dữ liệu, và nó được sử dụng rộng rãi trong nhiều các lĩnh vực, bao gồm lý do không chắc chắn [9,10], xác định mục tiêu [11], thiết kế bộ điều khiển [12,13], an toàn sản xuất công nghiệp [14], phân loại [15,16], v.v. [17]. Đồng thời, Tham khảo [18]đề xuất quy tắc kết hợp Mô hình niềm tin có thể chuyển đổi (Transferable Belief Model - TBM), đây là phiên bản không chuẩn hóa của quy tắc kết hợp Dempster (DCR). Cả hai quy tắc kết hợp đều có tính chất giao hoán và liên kết, và cả hai đều cho rằng các mục bằng chứng được kết hợp là khác biệt [19]. Tuy nhiên, khi dữ liệu bằng chứng được thu thập có tính xung đột cao, DCR thường không tạo ra kết quả trực quan. Để giải quyết vấn đề này, các học giả đã đề xuất nhiều giải pháp, trong đó đề xuất hàm niềm tin cơ sở [20] dựa trên quy mô khung nhận thức (FOD). Phương pháp này có thể loại bỏ xung đột cao trong bằng chứng bằng cách sửa đổi chỉ định xác suất cơ bản (BPA) trong FOD đầy đủ, do đó tạo ra kết quả trực quan. Nó phù hợp cho các hệ thống quân sự và thời gian thực khác cập nhật hệ thống [20]. Đồng thời, chức năng niềm tin cơ sở đã được sử dụng rộng rãi và mở rộng kể từ khi nó được đề xuất, chẳng hạn như trong [21,22]. Tuy nhiên, chức năng niềm tin cơ sở có một số hạn chế. Nó không xem xét thông tin không chắc chắn do FOD không đầy đủ gây ra, vì vậy nó chỉ có thể được sử dụng trong FOD đầy đủ.

Quy tắc kết hợp tổng quát (The generalized combination rule-GCR) trong FOD không đầy đủ có cùng một vấn đề như DCR cổ điển. Khi đối mặt với bằng chứng mâu thuẫn cao, nó thường tạo ra kết quả kết hợp điều đó trái ngược với trực giác. Hơn nữa, với giả định FOD không đầy đủ, các nguồn thông tin không chắc chắn phức tạp hơn. Trong số đó, thông tin không chắc chắn được đại diện bởi hàm khối lượng khác không của tập hợp rỗng và tính không đầy đủ có thể có của FOD [23] bị bỏ qua bởi hàm niềm tin cơ sở. Để giải quyết các vấn đề trên, bài báo này mở rộng cơ sở niềm tin chức năng và đề xuất một phương pháp để sửa đổi BPA trong FOD không đầy đủ. Phương pháp này không chỉ kế thừa các đặc điểm ban đầu của chức năng niềm tin cơ sở, mà còn tính đến giá trị của hàm khối lượng tập hợp rỗng khác không, cho phép sửa đổi BPA ở trạng thái mở FOD không cạn kiệt. Phương pháp này không chỉ có thể được áp dụng cho FOD không đầy đủ, mà còn có thể được giảm xuống chức năng niềm tin cơ sở trong FOD đầy đủ. Hơn nữa, chúng tôi còn thấy rằng entropy tăng đáng kể sau khi BPA được sửa đổi bằng phương pháp đề xuất. Niềm tin entropy được sử dụng để đo lường sự không chắc chắn của thông tin, có thể hiển thị xung đột một cách trực quan hơn. Sự gia tăng về giá trị của niềm tin entropy là hệ quả của xung đột. Bài báo này cũng đề xuất một dữ liệu xung đột phương pháp quản lý dựa trên hàm niềm tin cơ sở mở rộng (EBBF) và xác minh tính khả thi và hiệu quả của phương pháp đề xuất thông qua việc phân tích một số ví dụ. Các bước xung đột phương pháp quản lý dữ liệu như sau: Đầu tiên, chúng tôi tính toán giá trị của EBBF, sau đó sửa đổi BPA tương ứng với phương pháp sửa đổi được đề xuất và cuối cùng sử dụng DCR hoặc GCR cho dữ liệu tổng hợp để thu được kết quả hợp lý.

Phần còn lại của bài viết này là như sau:

+ Phần 2 giới thiệu các công việc liên quan đến nền kiến thức.

+ Phần 3 giới thiệu kiến thức sơ bộ.

+ Phần 4 đề xuất EBBF và một số các đặc tính của nó.

+ Phần 5 giới thiệu phương pháp quản lý xung đột dựa trên EBBF và đưa ra một số ví dụ và ứng dụng để kiểm chứng tính hiệu quả của phương pháp. Hơn nữa, trong phần 5, sự khác biệt giữa hai quy tắc kết hợp được đề cập ở trên cũng được so sánh và thảo luận.

+ Phần 6 rút ra kết luận của bài viết này.

1. **Công việc liên quan**

Trong lĩnh vực hợp nhất thông tin và quản lý dữ liệu xung đột, các học giả đã áp dụng nhiều phương pháp quản lý dữ liệu xung đột [24]. Hạng mục đầu tiên là xây dựng lại các quy tắc kết hợp. Nhiều nhà nghiên cứu cố gắng xây dựng lại quy tắc kết hợp để xử lý các dữ liệu có tính xung đột cao [23,25,26]. Trong số đó, Yager chỉ ra rằng cần phải loại bỏ yếu tố chuẩn hóa và đặt nó vào miền chưa biết để có được một quyết định hợp lý [27]. Sau đó, dựa trên ý tưởng của Yager, Dubois và Prade đã đề xuất một toán tử kết hợp cụ thể hơn [28]. Lefevre et al, phát triển một khuôn khổ chung để thống nhất một số quy tắc kết hợp cổ điển, có tính đến việc sử dụng đào tạo bộ và tối thiểu hóa hệ số trọng số của tiêu chí lỗi [29]. Su et al. đề xuất một quy tắc mới để kết hợp các phần của bằng chứng từ quan điểm về sự phụ thuộc và độc lập của bằng chứng [30].

Loại thứ hai là tiền xử lý chức năng khối lượng của bằng chứng [31,32]. Quá trình tiền xử lý BPA có thể loại bỏ hiệu quả các hiện tượng hoàn toàn mâu thuẫn giữa bằng chứng, có thể tránh được các kết quả hợp nhất không trực quan do xung đột bằng chứng gây ra [20]. Việc phủ định BPA cũng được đề xuất để giải quyết thông tin không chắc chắn trong bằng chứng [33,34]. Chiến lược trong bài báo này là xử lý trước BPA. Hiện nay, các phương pháp nghiên cứu chính của BPA tiền xử lý như sau.

* 1. Phương pháp tiền xử lý BPA dựa trên tập mờ

Trong trường hợp sử dụng phương pháp tạo tri thức trước cho các mẫu, Xu et al. được tạo ra một BPA lồng nhau sử dụng hàm mật độ xác suất [35]. Loại BPA này vượt trội hơn BPA với những người độc thân ở một mức độ nào đó vì nó ít gây ra xung đột hơn. Tuy nhiên, trên thực tế, nền công nghiệp hiện đại và công nghệ đang phát triển nhanh chóng. Do đó, rất có thể có một số điều chưa biết các loại phần tử bên ngoài FOD đã được thiết lập trước đó [23,36]. Để giải quyết loại vấn đề, Zhang và Deng đã đề xuất một phương pháp xác định điểm biên dựa trên tam giác số mờ [37]. Bằng cách xác định giá trị trung bình, độ lệch chuẩn, giá trị cực trị và tam giác chức năng thành viên của từng thuộc tính và sử dụng giao điểm của mẫu thử nghiệm và mẫu ở trên mô hình, một hàm BPA lồng nhau đã được xây dựng trong [37], có thể gán giá trị cho các tập hợp trống. Phương pháp này phù hợp với cả FOD đầy đủ và FOD không đầy đủ, yêu cầu ít dữ liệu trước hơn và được thúc đẩy bởi dữ liệu, vì vậy nó có thể làm giảm sự không chắc chắn chủ quan. Ngoài ra, nhiều học giả đã đề xuất phương pháp đo độ không đảm bảo dựa trên tập mờ, chẳng hạn như [38–41].

* 1. Phương pháp tiền xử lý BPA với Entropy niềm tin

Trong lĩnh vực tổng hợp thông tin, nhiều học giả đã đề xuất niềm tin entropy để đo lường thông tin không chắc chắn [42–46]. Có thể sử dụng hệ số trọng số được tính toán bởi niềm tin entropy để sửa đổi dữ liệu xung đột [47]. Shannon entropy [48] được áp dụng để đo lường thông tin không chắc chắn trong một khuôn khổ xác suất, và đã được công nhận rộng rãi và mở rộng cho nhiều lĩnh vực, chẳng hạn như entropy mạng trong các mạng phức tạp [49] và phân tích khuếch đại gen trong lĩnh vực tin sinh học [50]. Tuy nhiên, Shannon entropy không thể được áp dụng trực tiếp cho phép đo thông tin không chắc chắn của hàm khối lượng trong lý thuyết bằng chứng. Để giải quyết vấn đề này vấn đề, nhiều biện pháp không chắc chắn trong lý thuyết bằng chứng đã được đề xuất, bao gồm cả sự mơ hồ đo lường [51], Deng entropy [52], v.v. [53,54]. Deng entropy [52] xem xét sự không chắc chắn thông tin được mang bởi hàm khối lượng. Deng entropy không chỉ có thể suy biến thành Shannon entropy trong những điều kiện nhất định mà còn đưa ra phép đo hợp lý trong nhiều môi trường phức tạp [48]. Hơn nữa, Deng entropy đã được sử dụng trong nhiều ứng dụng thực tế, chẳng hạn như chẩn đoán lỗi [55], ra quyết định [56] và hợp nhất dữ liệu cảm biến [34,57]. Tuy nhiên, Deng entropy không tính đến tính đến quy mô của FOD, đây cũng là một nguồn thông tin không chắc chắn quan trọng. Sự thiếu thông tin như vậy có thể dẫn đến giảm hiệu quả xử lý thông tin và thậm chí không có khả năng đối phó với một số thông tin không chắc chắn một cách hiệu quả. Dựa trên kích thước của FOD, Tang et al. đã đề xuất Deng entropy có trọng số [58], làm mất tính nhất quán xác suất mà Deng entropy hài lòng, nhưng mô hình hóa thông tin không chắc chắn hơn trong cơ thể bằng chứng. Sau khi BPA được sửa đổi bằng phương pháp này, sự mất mát thông tin được giảm thiểu một cách hiệu quả và kết quả hợp nhất hợp lý hơn có thể được thu được. Tuy nhiên, Deng entropy không xác minh các thuộc tính cần thiết cho loại phép đo này và trình bày một số hành vi không mong muốn [59] trong một số trường hợp. Mặc dù một số học giả đã đề xuất một biến đổi Deng entropy [58,60], công trình trong [61] đã chứng minh rằng những biến đổi này vẫn không thể đáp ứng hầu hết các tính chất toán học cần thiết và chúng đã trình bày hầu hết các khiếm khuyết về hành vi tồn tại trong Deng entropy. Do đó, Deng entropy và sửa đổi tương ứng phải là thận trọng trong các ứng dụng thực tế.

* 1. Phương pháp tiền xử lý BPA sử dụng hàm niềm tin cơ sở

Đối với phương pháp xây dựng lại luật kết hợp, Haenni đã đề xuất trong bài báo [62] rằng khi có quá nhiều bằng chứng, việc gán cho từng yếu tố trọng lượng là không thực tế và khi có nhiều tập hợp con trong FOD, số tiền tính toán tăng theo cấp số nhân. Để sửa đổi chức năng khối lượng, nó cần ghi lại lượng dữ liệu và tính toán mức độ tương tự hoặc tương quan của dữ liệu, điều này làm tăng thời gian tính toán nên khó thực hiện trong trường hợp yêu cầu thời gian thực cao. trong tầm nhìn của giới hạn trên, Tài liệu tham khảo [20] đề xuất một hàm niềm tin cơ sở mới. Phương pháp này duy trì các đặc điểm tốt của DCR và độ phức tạp tính toán thấp. Hơn nữa, phương pháp này có thể loại bỏ sự mâu thuẫn hoàn toàn giữa các bằng chứng. Dựa trên hàm niềm tin cơ sở, nhiều phương pháp đã được đề xuất bởi các học giả khác, chẳng hạn như của [22,63], có thể giải quyết vấn đề mà DCR không thể thu được kết quả trực quan khi áp dụng cho dữ liệu có tính xung đột cao.Tuy nhiên, các phương pháp trên không phù hợp với FOD không triệt để. Phương pháp đề xuất trong bài báo này là một phần mở rộng của hàm niềm tin cơ sở để làm cho nó có thể áp dụng cho FOD không đầy đủ.

1. **Sơ bộ**

Một số sơ bộ được giới thiệu ngắn gọn trong phần này, bao gồm Dempster–Shafer (D-S) lý thuyết bằng chứng [5,6], hàm niềm tin cơ sở [20], DCR [5,6], quy tắc kết hợp tổng quát [23] v..v...

* 1. Lý thuyết bằng chứng Dempster–Shafer cổ điển

**Định nghĩa 1:** Khung phân biệt Ω được định nghĩa là một tập hữu hạn không rỗng chứa N loại trừ lẫn nhau sự kiện và biểu hiện cụ thể của nó như sau:

Ω = { Θ1, Θ2, ..., Θi , ..., ΘN } (1)

**Định nghĩa 2:** Cho Ω, phép gán xác suất cơ bản (BPA) (hay hàm khối lượng) là ánh xạ m: 2Ω → [0, 1], đáp ứng các thuộc tính sau:

m(∅) = 0, = 1 (2)

Nếu m(A) > 0 thì tập con A được gọi là phần tử tiêu điểm và m(A) > 0 là giá trị hàm khối lượng của mệnh đề tập con A

**Định nghĩa 3:** Trong lý thuyết bằng chứng D-S, DCR có thể hợp nhất hai hàm khối độc lập, m1 và m2:

m(A) = (m1 ⊕ m2)(A) = (3)

trong đó k là hệ số chuẩn hóa được định nghĩa như sau:

k = (4)

Điều đáng chú ý là các định nghĩa cổ điển của lý thuyết bằng chứng D-S được định nghĩa và sử dụng trong FOD đầy đủ

**Định nghĩa 4:** Trong giả thuyết FOD không đầy đủ, DCR được mở rộng bởi Deng trong [52]. Giao của tập rỗng và tập rỗng vẫn là tập rỗng thỏa mãn điều kiện ∅1 ∩ ∅2 = ∅. Cho hai BPAs (m1 và m2), quy tắc kết hợp tổng quát (GCR) được định nghĩa như sau:

m (A) = (5)

K = , (6)

m(∅) = m1(∅) · m2(∅), (7)

m(∅) = 1 khi và chỉ khi K = 1 (8)

* 1. Bình thường hóa và Quy tắc kết hợp TBM

**Định nghĩa 5:** Một BPA m được cho là không chuẩn nếu ∅ là một tập hợp tiêu cự (m(∅) 6= 0), có thể được biến đổi bởi phép toán chuẩn hóa được định nghĩa như sau [19]:

m(A) = k ∗ m(A) khi và chỉ khi A # ∅ (9)

thì m(A) = 0 (10)

Cho tất cả A ⊆ Ω, với k = (1 − m(∅))−1 .

**Định nghĩa 6:** Quy tắc kết hợp TBM được ghi chú với ?; giả sử m1 và m2 là hai BPA, và để m1 và m2 là kết quả của sự kết hợp bởi , như sau [19]:

m12(A) = (11)

Cho tất cả A ⊆ Ω

* 1. Hàm niềm tin cơ bản

**Định nghĩa 7:** Là một công cụ quản lý xung đột dữ liệu, hàm niềm tin cơ sở được đề xuất trong [20] và hàm niềm tin cơ sở được định nghĩa là:

mb (Ai) = (12)

trong đó Ai là một tập con tùy ý của Ω, N là kích thước của FOD, và sau đó BPA ban đầu được sửa đổi theo công thức sau:

m‘(Ai) = (13)

Chức năng niềm tin cơ sở có thể loại bỏ hiệu quả xung đột tuyệt đối giữa các dữ liệu và phù hợp với các hệ thống cập nhật thời gian thực lớn, chẳng hạn như hệ thống quân sự.

* 1. Sự mở rộng của Deng Entropy trong giả định FOD không đầy đủ

**Định nghĩa 8:** Phần mở rộng của Deng entropy [52] là một phương pháp đo entropy, kéo dài từ FOD toàn diện đến FOD không đầy đủ, và định nghĩa của nó như sau [64]:

Eebeow = - (14)

trong đó |A| đại diện cho số lượng phần tử chứa trong mệnh đề A, X đại diện cho FOD và | X| đại diện cho tiềm năng của FOD, đại diện cho số lượng các yếu tố được biết đến chính xác trong FOD. là một hàm giới hạn trên (hàm CEILING), liên quan đến việc làm tròn biến độc lập, nghĩa là thành một số nguyên không nhỏ hơn biến độc lập; Ví dụ: 0,2 = 1.

1. **Hàm niềm tin cơ sở mở rộng**

Trong phần này, chúng tôi đề xuất một hàm phân phối niềm tin trong FOD không đầy đủ dựa trên hàm niềm tin cơ sở, sử dụng giá trị của hằng số tự nhiên e và hàm khối lượng tập rỗng như một yếu tố đặc trưng mà hàm này được sử dụng để sửa đổi BPA. Trong khi đó, phương pháp này cũng được áp dụng cho FOD toàn diện

Người ta cho rằng FOD A chứa N phần tử loại trừ lẫn nhau; khi đó, A có 2 tập con N và trong FOD không đầy đủ, giá trị của hàm khối lượng tập rỗng không phải là 0. Dựa trên các giả định trên, hàm gán niềm tin cơ bản mở rộng được định nghĩa như sau:

neb(Ri) = (15)

trong đó Ri đại diện cho một tập con tùy ý của FOD, N là kích thước của FOD và m(∅) là giá trị hàm khối lượng của tập rỗng.

Sử dụng hàm phân phối niềm tin cơ sở mở rộng neb (Ri), BPA của mỗi bằng chứng được sửa đổi và phương pháp sửa đổi được định nghĩa như sau:

m’(Ri) = (16)

trong đó m(Ri) là giá trị của BPA của tất cả các bằng chứng. là một hàm giới hạn trên (hàm CEILING), liên quan đến việc làm tròn biến độc lập, nghĩa là thành một số nguyên không nhỏ hơn biến độc lập;Ví dụ: 0,2 = 1.

Mục đích của phép tính neb(Ri) là gán xác suất bằng nhau cho mỗi tập con của FOD trước khi BPA được tạo ra. Trước khi mọi nguồn bằng chứng xuất hiện, niềm tin của chúng ta vào mọi tình huống phải bình đẳng. neb (Ri) tương đương với những khả năng ban đầu này, có thể loại bỏ hiện tượng mâu thuẫn hoàn toàn giữa các bằng chứng. Phương pháp này phủ nhận tính tuyệt đối của bằng chứng, nghĩa là, ngay cả khi tất cả các bằng chứng hiện có đều chống lại một tập hợp con A, nhưng chúng tôi chưa tìm thấy tất cả các bằng chứng, vẫn có khả năng A là đúng. Đồng thời, giá trị khối lượng nonzero của tập rỗng được thêm vào neb (Ri) và quá trình hợp nhất làm hệ số đặc trưng. Nghĩa là, trong FOD không đầy đủ, sự không chắc chắn thông tin gây ra bởi hàm khối lượng tập rỗng nonzero được tính đến trong công thức để phương pháp có thể được sử dụng trong FOD không đầy đủ và BPA sửa đổi vẫn đáp ứng tính chất cơ bản là tổng các giá trị hàm khối lượng của tất cả các tập con là 1.

1. **Áp dụng hàm niềm tin cơ sở mở rộng**

Trong phần này, chúng tôi đề xuất một phương pháp quản lý dữ liệu xung đột dựa trên EBBF và đưa ra một số ví dụ và ứng dụng liên quan để xác minh tính hợp lý và hiệu quả của nó.

* 1. Phương pháp quản lý dữ liệu xung đột dựa trên EBBF

Phần này đề xuất một phương pháp quản lý dữ liệu xung đột dựa trên EBBF để xác minh khả năng ứng dụng và hiệu quả của EBBF trong lĩnh vực hợp nhất thông tin. Hình 1 thiết kế khuôn khổ của phương pháp quản lý dữ liệu xung đột dựa trên EBBF và các bước chi tiết như sau:

Bước 1: Trong FOD không đầy đủ, có rất nhiều thông tin không chắc chắn trong các ứng dụng thực tế. Để mô hình hóa thông tin không chắc chắn một cách có hệ thống, trong khuôn khổ của lý thuyết bằng chứng Dempster-Shafer, bước đầu tiên là sử dụng BPA để mô hình hóa thông tin không chắc chắn.

Bước 2: Tính toán giá trị của EBBF: Khi dữ liệu có nhiều xung đột, cần có một phương pháp hợp lý và hiệu quả để xử lý trước BPA trước khi xử lý dữ liệu tiếp theo. Phương pháp này sử dụng EBBF để sửa đổi BPA. Công thức tính tương ứng của hàm niềm tin cơ sở mở rộng (BBF) như sau:

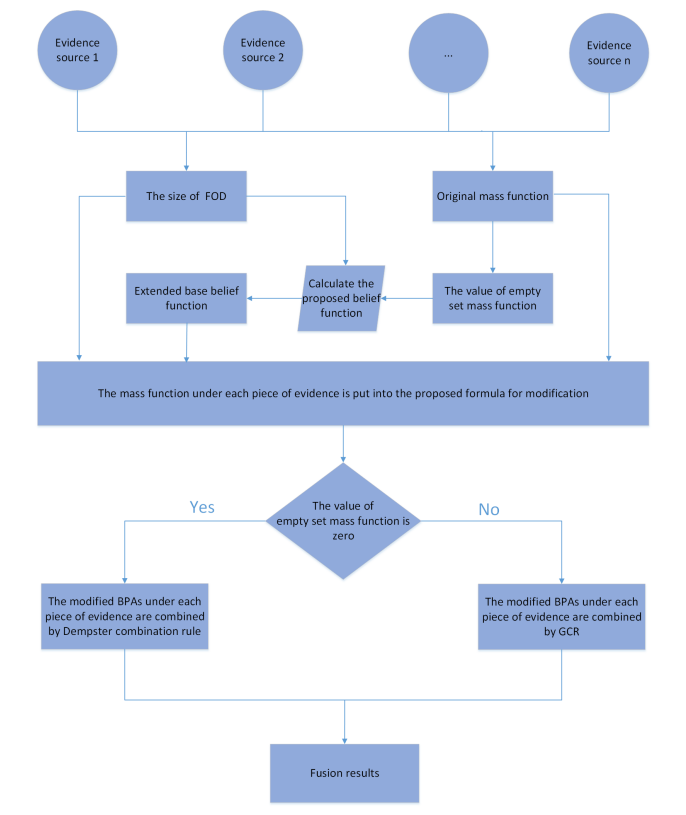
neb(Ri) = (17)

Bước 3: Theo giá trị EBBF được tính toán, BPA của mỗi nhóm được sửa đổi và công thức sửa đổi như sau:

m’(Ri) = (18)

Bước 4: Theo đó hàm khối lượng tập rỗng là 0, DCR hay GCR được chọn để hợp nhất dữ liệu. Theo DCR (n − 1), kết quả kết hợp của mỗi mệnh đề có thể thu được như sau:

m(A) = (((m1(A) ⊕ m2(A)) ⊕ m3(A))... ⊕ mn−1(A)) ⊕ mn(A), n ≥ 2 (19)



Hình 1. Khung phương pháp quản lý dữ liệu xung đột dựa trên hàm niềm tin cơ sở mở rộng (EBBF)

* 1. Ví dụ minh họa

Ví dụ 1: Trong ví dụ cổ điển rằng FOD là X = {a, b, c}, hai BPA được cho là:

m1({a}) = 0.99, m1({c}) = 0.01, m1({∅}) = 0 (20)

m2({b}) = 0.99, m2({c}) = 0.01, m2({∅}) = 0. (21)

Trong ví dụ này, hàm khối lượng của tập rỗng là 0, chỉ ra rằng hàm khối lượng được phân bổ trong FOD toàn diện. Giá trị EBBF được tính như sau:

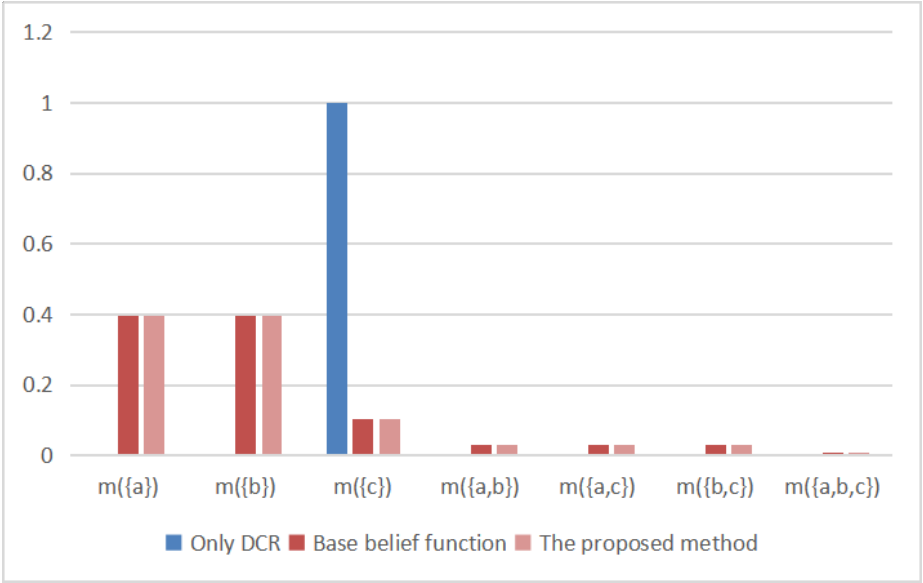
neb(Ri) = = 0.1429 = mb (Ai) (22)

Nó hoàn toàn giống với giá trị được tính với hàm niềm tin cơ bản. BPA của hai nhóm được sửa đổi tương ứng và kết quả được hiển thị trong Bảng 1.

Bảng 1. Chỉ định xác suất cơ bản đã sửa đổi (BPA) với phương pháp được đề xuất trong Ví dụ 1

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| BPA | m({a}) | m({b}) | m({c}) | m({a,b}) | m({a,c}) | m({b,c}) | m({a,b,c}) | m({∅}) |
| m1(Ri) | 0.5664 | 0.0714 | 0.0764 | 0.0714 | 0.0714 | 0.0714 | 0.0714 | 0.0714 |
| m2(Ri) | 0.0714 | 0.5664 | 0.0764 | 0.0714 | 0.0714 | 0.0714 | 0.0714 | 0.0714 |

Cuối cùng, kết quả của việc hợp nhất dữ liệu được thể hiện trong Hình 2. Sau khi xem xét và nghiên cứu, phương pháp hợp nhất dữ liệu mà chúng tôi áp dụng ở đây là DCR thay vì quy tắc liên hợp TBM. Theo công thức của hai phương pháp, sự khác biệt giữa hai phương pháp có thể được phân tích như sau: Cả quy tắc liên hợp TBM và DCR chủ yếu xem xét ảnh hưởng của giao điểm, nhưng quy tắc liên hợp TBM không được chuẩn hóa, điều này làm cho tổng các giá trị niềm tin ít hơn hơn 1. Rõ ràng, kết quả như vậy là phản trực giác, vì vậy quy tắc liên từ TBM không phù hợp với các ví dụ trong bài viết này.



Hình 2. Kết quả hợp nhất của hai phương pháp trong Ví dụ 2

Kết quả hợp nhất cho thấy rằng quy tắc DCR không thể tạo ra kết quả logic trong ví dụ này, nhưng cả phương pháp trong bài báo này và phương pháp hàm niềm tin cơ sở đều có thể tạo ra kết quả trực quan. Cũng có thể thấy rằng với giả định FOD đầy đủ, phương pháp trong bài viết này có thể được rút gọn thành hàm niềm tin cơ sở.

Ví dụ 2: Giả sử rằng FOD là X = {a, b}, và hai BPA được cho là:

m1({a}) = 0.9, m1({∅}) = 0.1 (23)

m2({a}) = 0.9, m2({∅}) = 0.1 (24)

Tại thời điểm này, m(∅) 6= 0, cho thấy FOD không đầy đủ và ví dụ này nằm trong giả định FOD không đầy đủ. Quá trình tính toán được thể hiện trong H1

Bước 1: Từ FOD X = {a, b}, chúng ta có thể biết rằng tiềm năng của FOD là: N = 2

Bước 2: Đưa giá trị hàm khối lượng tập rỗng và N vào công thức và tính giá trị EBBF như sau:

neb(Ri) = = 0.3454 (25)

Bước 3: Sử dụng giá trị của hàm niềm tin để sửa đổi BPA của từng mẩu bằng chứng và tính toán như sau:

m1({a}) = = 0.5229 (26)

m1({b}) = m1({a,b}) = = 0.1450 (27)

m1({∅}) = = 0.1870 (28)

m2({a}) = = 0.5229 (29)

m2({b}) = m2({a,b}) = = 0.1450 (30)

m2({∅}) = = 0.1870 (31)

Bước 4: Đánh giá xem hàm khối lượng tập hợp rỗng có bằng 0 hay không và chọn một quy tắc kết hợp khác.

Trong ví dụ này, giá trị hàm khối lượng của tập hợp trống không phải là 0, vì vậy GCR được sử dụng để hợp nhất các BPA của hai tập hợp bằng chứng và thu được các kết quả sau:

m({a}) = 0.8053, m({b}) = 0.1195 (32)

m({a, b}) = 0.0398, m({∅}) = 0.0350 (33)

Kết quả cho thấy sự ủng hộ cao đối với mệnh đề a, phù hợp với tình hình thực tế của bằng chứng. Tuy nhiên, đồng thời, kết quả cũng cho b một mức độ hỗ trợ nhỏ nhất định, điều này cho thấy rằng chúng tôi chưa thu thập được tất cả các bằng chứng; vẫn có khả năng b đúng, tránh số liệu quá tuyệt đối.

Ví dụ 3: Trong FOD X = {a, b}, có hai bộ BPA xung đột cao, như sau:

m1({a}) = 0.9, m1({∅}) = 0.1 (34)

m2({b}) = 0.9, m2({∅}) = 0.1 (35)

Sử dụng phương pháp được đề xuất để tính giá trị hàm niềm tin, chúng ta có thể nhận được:

neb(Ri) = 0.3454 (36)

Sau đó, neb(Ri) được sử dụng để sửa đổi giá trị BPA của mỗi bằng chứng, như thể hiện trong Bảng 2.

Bảng 2. BPA được sửa đổi với phương pháp được đề xuất trong Ví dụ 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| BPA | m({a}) | m({b}) | m({a,b}) | m({∅}) |
| m1(Ri) | 0.5229 | 0.1450 | 0.1450 | 0.1870 |
| m2(Ri) | 0.1450 | 0.5229 | 0.1450 | 0.1870 |

Các kết quả hợp nhất thu được bởi GCR được thể hiện trong Bảng 3. Dữ liệu cho thấy phương pháp được đề xuất đưa ra cùng một mức độ hỗ trợ cho a và b trong hai bộ dữ liệu hoàn toàn mâu thuẫn, đây là câu trả lời chính xác trong trực giác. Nó đã được chứng minh rằng phương pháp được đề xuất vẫn có thể cho kết quả tốt phù hợp với các sự kiện trong dữ liệu mâu thuẫn cao.

Bảng 3. Kết quả hợp nhất trong Ví dụ 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| BPA | m({a}) | m({b}) | m({a,b}) | m({∅}) |
| Phương pháp đề xuất | 0.4548 | 0.4548 | 0.0554 | 0.0350 |

Ví dụ 4: Trong FOD X = {a, b, c}, các BPA như sau:

m1({a}) = 0.9, m1({a, b}) = 0.01, m1({∅}) = 0.09 (37)

m2({b}) = 0.01, m2({c}) = 0.91, m2({∅}) = 0.08. (38)

Sử dụng phương pháp được đề xuất để tính toán các giá trị của EBBF, chúng ta có thể nhận được:

m1 : neb(Ri)1 = 0.1448, m2 : neb(Ri)2 = 0.1446 (39)

Theo giá trị của hai EBBF, BPA của hai nhóm đã được sửa đổi cho phù hợp và kết quả được thể hiện trong Bảng 4

**Bảng 4.** BPA được sửa đổi bằng phương pháp được đề xuất trong Ví dụ 4

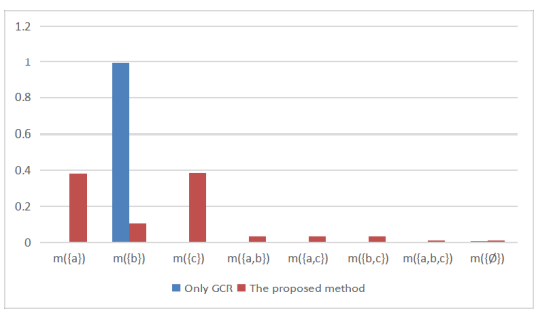
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **BPA** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| m1(Ri) | 0.4841 | 0. 0671 | 0. 0671 | 0. 0717 | 0. 0671 | 0. 0671 | 0.0671 | 0. 1088 |
| m2(Ri) | 0. 0670 | 0. 0717 | 0. 4890 | 0. 0670 | 0. 0670 | 0. 0670 | 0.0670 | 0. 1041 |

Sau đó, GCR được sử dụng cho phản ứng tổng hợp để thu được kết quả. Đồng thời, BPA từ bằng chứng trong bảng được hợp nhất trực tiếp với GCR và kết quả của hai phương pháp là được so sánh, như thể hiện trong Bảng 5 và Hình 3.

Trong trường hợp này, hai nhóm bằng chứng tương ứng ủng hộ cao cho a và c, nhưng kết quả hợp nhất chỉ sử dụng GCR sẽ gán mức hỗ trợ cao nhất cho đề xuất b, điều này rõ ràng là không hợp lý. Tuy nhiên, sau khi BPA được sửa đổi bằng phương pháp được đề xuất trong bài báo này, kết quả phản ứng tổng hợp

thu được bởi GCR đã hỗ trợ rất nhiều cho cả a và c. Hơn nữa, như là sự hỗ trợ cho c trong mảnh bằng chứng thứ hai lớn hơn một chút so với mảnh bằng chứng đầu tiên, kết quả hợp nhất cũng mang lại một chút lợi thế cho c. Điều này chỉ ra rằng trong trường hợp dữ liệu có tính xung đột cao, quyền câu trả lời trực quan không thể thu được đơn giản bằng cách sử dụng GCR, trong khi kết quả hợp lý và trực quan có thể thu được bằng cách sử dụng phương pháp trong bài viết này.

Các ví dụ trên xác minh rằng phương pháp được đề xuất tương thích với hàm niềm tin cơ sở trong FOD đầy đủ và xác minh tính khả thi và hiệu quả của phương pháp được đề xuất. Sau đây các ví dụ thảo luận về một số thuộc tính khác của phương pháp được đề xuất.



**Hình 3.** So sánh kết quả hợp nhất trong Ví dụ 4.

**Bảng 5.** Kết quả của hai phương pháp kết hợp trong Ví dụ 4.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Fusion Methods** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| Only GCR | 0.0000 | 0.0660 | 0. 0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0. 0072 |
| The proposed method | 0.3831 | 0.1061 | 0.3851 | 0. 0354 | 0. 0338 | 0. 0338 | 0.0113 | 0.0113 |

**Ví dụ 5.** Trong FOD X = {a, b, c}, các BPA như sau:

m1({a}) = 0.7, m1({b}) = 0.1, m1({c}) = 0.1, m1({∅}) = 0.1 (40)

m2({a}) = 0.1, m2({b}) = 0.1, m2({c}) = 0.7, m2({∅}) = 0.1. (41)

Sử dụng phương pháp được đề xuất để tính giá trị của EBBF, chúng ta có thể nhận được:

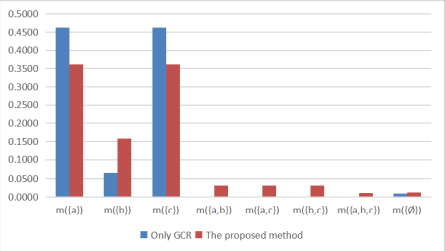
neb(Ri)1 = neb(Ri)2 = = 0.1454. (42)

Sau đó, BPA của từng mẩu bằng chứng đã được sửa đổi bằng cách sử dụng giá trị của hàm niềm tin được tính toán, và kết quả được thể hiện trong Bảng 6.

**Bảng 6.** BPA được sửa đổi bằng phương pháp được đề xuất trong Ví dụ 5

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **BPA** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| m1(Ri) | 0.3912 | 0. 1134 | 0. 1134 | 0. 0671 | 0. 0671 | 0. 0671 | 0.0671 | 0. 1134 |
| m2(Ri) | 0. 1134 | 0. 1134 | 0. 3912 | 0. 0671 | 0. 0671 | 0. 0671 | 0. 0671 | 0. 1134 |

Sau đó, GCR được sử dụng cho phản ứng tổng hợp để thu được kết quả. Đồng thời, BPA từ bằng chứng trong bảng được hợp nhất trực tiếp với GCR và kết quả của hai phương pháp là được so sánh, như thể hiện trong Bảng 7 và Hình 4



**Hình 4.** So sánh kết quả hợp nhất trong Ví dụ 5.

**Bảng 7.** Kết quả của hai phương pháp kết hợp trong Ví dụ 5.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Fusion Methods** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| Only GCR | 0.4620 | 0.0660 | 0. 4620 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0100 |
| The proposed method | 0.3620 | 0.1578 | 0.3620 | 0. 0316 | 0. 0316 | 0.0316 | 0.0105 | 0.0129 |

Theo Bảng 7, khi hàm khối lượng của từng tập hợp con phần tử đơn lẻ trong FOD không bằng 0, kết quả tổng hợp thu được bằng cách sử dụng phương pháp trong bài viết này để sửa đổi BPA và chỉ sử dụng GCR đều phản ánh rằng a và c có độ tin cậy cao như nhau. Chúng tôi đã tính thêm entropy các giá trị của hàm khối lượng bằng chứng ban đầu và hàm khối lượng được sửa đổi. Ở đây, chúng tôi đã thông qua mở rộng Deng entropy (EBEOW) [64], có thể tính toán giá trị entropy ở dạng không đầy đủ FOD. Theo tính toán, giá trị entropy Deng mở rộng của các BPA ban đầu là:

Eebeow(m)1 = 2.9237. (43)

Entropy Deng mở rộng của các BPA được sửa đổi bằng phương pháp đề xuất trong bài báo này là như sau:

Eebeow(m)2 = 3.8089. (44)

Rõ ràng, entropy tương ứng với kết quả tổng hợp của phương pháp được đề xuất tăng lên đáng kể, và sự gia tăng của entropy cho thấy rằng việc gán niềm tin bị phân tán hơn và sự không chắc chắn được tăng lên, đó là hậu quả của cuộc xung đột. Nói chung, khi đề xuất của một tập hợp con phần tử duy nhất khác không và dữ liệu rất mâu thuẫn, phương pháp được đề xuất trong phần này bài báo gán một phần niềm tin cho các phần tử đa tập hợp con khác, do đó giảm thiểu rủi ro.

**Ví dụ 6.** Giả sử FOD là X = {a, b, c} và hai BPA được cho là:

m1({a}) = 0.9, m1({a, b, c}) = 0.05, m1({∅}) = 0.05 (45)

m2({b}) = 0.9, m2({a, b, c}) = 0.05, m2({∅}) = 0.05 (46)

Theo phương pháp đề xuất, giá trị của hàm niềm tin tương ứng với bằng chứng m1 và m2 là được tính như sau:

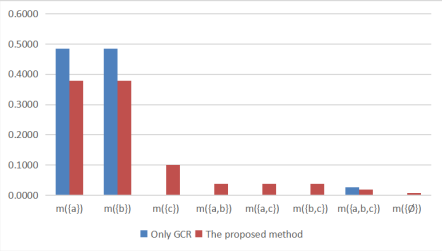
neb(Ri)1 = neb(Ri)2 = 0.1439 (47)

Sau đó, BPA của từng mẩu bằng chứng đã được sửa đổi bằng cách sử dụng giá trị của hàm niềm tin được tính toán, và kết quả được thể hiện trong Bảng 8.

**Bảng 8.** BPA được sửa đổi bằng phương pháp được đề xuất trong Ví dụ 6.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **BPA** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| m1(Ri) | 0.4852 | 0. 0669 | 0. 0669 | 0. 0669 | 0. 0669 | 0. 0669 | 0.0901 | 0. 0901 |
| m2(Ri) | 0.0669 | 0. 4852 | 0. 0669 | 0. 0669 | 0. 0669 | 0. 0669 | 0. 0901 | 0. 0901 |

Cuối cùng, kết quả hợp nhất thu được với GCR và kết quả hợp nhất với không sửa đổi BPA mà GCR thu được được trình bày trong Bảng 9 và Hình 5.



**Hình 5.** So sánh kết quả hợp nhất trong Ví dụ 6.

**Bảng 9.** Kết quả của hai phương pháp kết hợp trong Ví dụ 6.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Fusion Methods** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| Only GCR | 0.4853 | 0. 4853 | 0. 0000 | 0.0000 | 0.0000 | 0.0000 | 0.0270 | 0.0025 |
| The proposed method | 0.3796 | 0. 3796 | 0.0998 | 0.0380 | 0. 0380 | 0. 0380 | 0.0187 | 0.0081 |

Có thể thấy từ các kết quả trong Hình 5 rằng khi tất cả các hàm khối lượng của một tập hợp hoàn chỉnh là khác không, kết quả hợp nhất thu được bằng cách sử dụng phương pháp sửa đổi BPA được đề xuất trước và bằng cách chỉ sử dụng GCR, cả hai đều phản ánh rằng a và b có niềm tin cao như nhau, đây là một trực giác kết quả. Entropies niềm tin mở rộng của hàm khối lượng bằng chứng ban đầu và khối lượng được sửa đổi chức năng được tính toán thêm. Theo tính toán, entropy niềm tin mở rộng của BPA là Eebeow(m)1 = 2,8185 và entropy niềm tin mở rộng của BPA được sửa đổi theo phương pháp đề xuất là Eebeow(m)2 = 3,8488. Giá trị entropy tương ứng với kết quả hợp nhất của phương pháp này rõ ràng tăng. Do đó, khi hàm khối lượng của toàn bộ tập con phần tử của mỗi bằng chứng không bằng 0 và dữ liệu rất mâu thuẫn, phương pháp được đề xuất sẽ phân bổ một phần mức độ niềm tin cho những người khác đề xuất phụ, do đó làm giảm rủi ro.

**Ví dụ 7.** Trong FOD X = {a, b, c}, các BPA như sau:

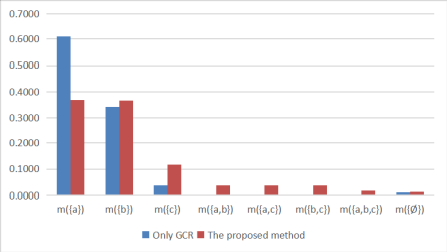
m1({a}) = 0.8, m1({b}) = 0.05, m1({c}) = 0.05, m1({∅}) = 0.1 (48)

m2({b}) = 0.8, m2({a, b, c}) = 0.1, m2({∅}) = 0.1. (49)

Sử dụng phương pháp được đề xuất để tính giá trị hàm niềm tin, chúng ta có thể nhận được:

neb(Ri)1 = neb(Ri)2 = 0.1450. (50)

Sau đó, GCR được sử dụng cho phản ứng tổng hợp để thu được kết quả. Đồng thời, BPA từ bằng chứng trong bảng được hợp nhất trực tiếp với GCR và kết quả của hai phương pháp là được so sánh, như thể hiện trong Bảng 10 và Hình 6.



**Hình 6.** So sánh kết quả hợp nhất trong Ví dụ 7.

**Bảng 10.** Kết quả của hai phương pháp kết hợp trong Ví dụ 7.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Fusion Methods** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| Only GCR | 0.6092 | 0.3427 | 0.0381 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0100 |
| The proposed method | 0.3699 | 0.3669 | 0.1176 | 0.0384 | 0.0384 | 0.0384 | 0.0176 | 0.0129 |

Khi hàm khối lượng tập hợp con đầy đủ của một tập hợp bằng chứng khác không, giá trị của tất cả các hàm khối lượng tập hợp con phần tử đơn của tập hợp bằng chứng khác là khác không và dữ liệu rất cao mâu thuẫn, kết quả hợp nhất chỉ sử dụng GCR chỉ mang lại mức độ tin cậy cao, trong khi kết quả nhiệt hạch sử dụng phương pháp trong bài báo này cho a và b độ hỗ trợ cao. Rõ ràng, kết quả

của phương pháp đề xuất phù hợp hơn với thực tế khách quan.

* 1. Ứng dụng cho dữ liệu nhân tạo

Để xác minh tính khả dụng và hiệu quả của phương pháp quản lý dữ liệu xung đột được đề xuất trong phần này, một ví dụ trong [65] được sử dụng để phân tích ví dụ và kết quả tính toán được so sánh với các phương pháp khác.

Xem xét vấn đề nhận dạng mục tiêu, người ta cho rằng có ba mục tiêu tiềm năng là được biểu diễn lần lượt là a, b và c. Theo phương pháp quản lý dữ liệu xung đột dựa trên EBBF đề xuất trong phần này, bước đầu tiên là lập mô hình bằng chứng thông tin không chắc chắn. Các báo cáo của năm cảm biến được mô hình hóa bởi BPA, và kết quả được thể hiện trong Bảng 11. Bảng 11 sửa đổi một chút

dữ liệu trong [65], gây ra sự mở rộng từ FOD đầy đủ sang FOD không đầy đủ. phân tích trực quan nhận thấy rằng báo cáo từ cảm biến thứ hai không nhất quán và xung đột với bốn cảm biến khác báo cáo. Ngoài ra, bốn cảm biến khác báo cáo rằng a có nhiều khả năng là mục tiêu tiềm năng nhất bởi vì giá trị hàm khối lượng của nó là lớn nhất và độ tin cậy là cao nhất.

**Table 11.** BPAs of the application example.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **BPA** | **m({a})** | **m({b})** | **m({c})** | **m({a,c})** | **m({∅})** |
| 1st sensor report: m1(·) | 0.41 | 0.29 | 0.2 | 0 | 0.1 |
| 2nd sensor report: m2(·) | 0 | 0.9 | 0.05 | 0 | 0.05 |
| 3rd sensor report: m3(·) | 0.58 | 0.07 | 0 | 0.15 | 0.2 |
| 4th sensor report: m4(·) | 0.55 | 0.1 | 0 | 0.15 | 0.2 |
| 5th sensor report: m5(·) | 0.6 | 0.1 | 0.0 | 0.2 | 0.1 |

Dựa trên dữ liệu trong Bảng 11, các giá trị EBBF của từng nhóm bằng chứng đã được tính toán như sau:

neb(Ri)1 = 0.1450 (51)

neb(Ri)2 = 0.1439 (52)

neb(Ri)3 = 0.1475 (53)

neb(Ri)4 = 0.1475 (54)

neb(Ri)5 = 0.1450 (55)

Theo các giá trị EBBF, các BPA đã được sửa đổi và dữ liệu đã sửa đổi của tất cả các BPA trong Bảng 11 được thể hiện trong Bảng 12.

**Bảng 12.** BPAs trong ứng dụng sau khi sửa đổi với phương pháp đề xuất.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| BPA | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| m1(·) | 0.2569 | 0.2041 | 0.1597 | 0.0671 | 0.0671 | 0.0671 | 0.0671 | 0.1134 |
| m2(·) | 0.0669 | 0.4852 | 0.0901 | 0.0669 | 0.0669 | 0.0669 | 0.0669 | 0.0901 |
| m3(·) | 0.3337 | 0.0998 | 0.0677 | 0.0677 | 0.1365 | 0.0677 | 0.0677 | 0.1594 |
| m4(·) | 0.3199 | 0.1135 | 0.0677 | 0.0677 | 0.1365 | 0.0677 | 0.0677 | 0.1594 |
| m5(·) | 0.3449 | 0.1134 | 0.0671 | 0.0671 | 0.1597 | 0.0671 | 0.0671 | 0.1134 |

Hợp nhất dữ liệu bằng cách sử dụng quy tắc kết hợp tổng quát như sau:

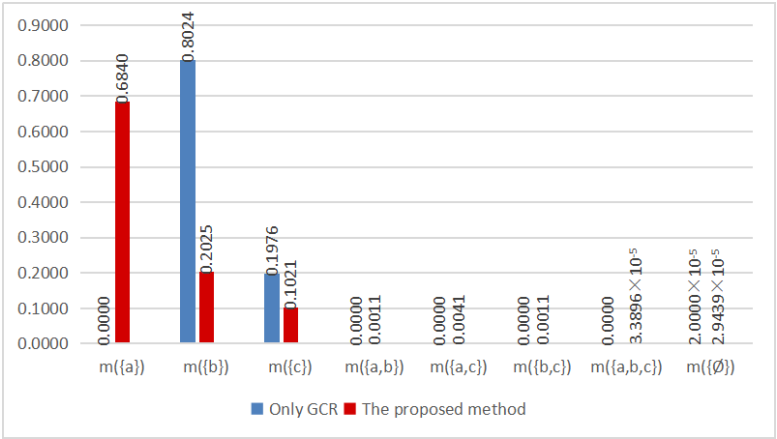
m(A) = (((m1(A) ⊕ m2(A)) ⊕ m3(A)) ⊕ m4(A)) ⊕ m5(A) = 0.6840. (56)

Sau đó, GCR được sử dụng cho phản ứng tổng hợp để thu được kết quả. Đồng thời, các BPA từ bằng chứng trong bảng được hợp nhất trực tiếp với GCR và kết quả của hai phương pháp là được so sánh, như thể hiện trong Bảng 13 và Hình 7.

Kết quả tổng hợp cho thấy rằng sử dụng phương pháp được đề xuất trong phần này, có thể kết luận rằng a là mục tiêu đã xác định, phù hợp với kết quả phân tích trực quan. Tuy nhiên, kết quả của phản ứng tổng hợp GCR mà không có EBBF cho thấy b là mục tiêu đã xác định, trong khi khả năng a là mục tiêu là 0, đây rõ ràng không phải là kết quả trực quan. Qua so sánh trên, ta có thể thấy rõ rằng các kết quả hợp nhất dữ liệu xác minh tính hiệu quả của phương pháp được đề xuất trong phần này, có thể được sử dụng tốt trong quản lý dữ liệu xung đột trong kỹ thuật thực tế.

**Bảng 13.** Kết quả của hai phương pháp kết hợp trong ứng dụng dữ liệu nhân tạo.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Fusion Methods** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| Only GCR | 0.0000 | 0.8024 | 0.1976 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 2\*10-5 |
| The proposed method | 0.6840 | 0.2025 | 0.1021 | 0.0011 | 0.0041 | 0.0011 | 3.3896\*10-5 | 2.9439\*10-5 |



**Hình 7.** So sánh kết quả hợp nhất của ứng dụng.

* 1. Ứng dụng để phân loại

Để xác minh tính hiệu quả của phương pháp quản lý dữ liệu xung đột được đề xuất, một bộ dữ liệu thực để phân loại từ Kho lưu trữ máy học của Đại học California Irvine (UCI) đã được áp dụng trong ứng dụng này. Trong bộ dữ liệu mống mắt của UCI, có ba loại tròng đen: Setosa (a), Versicolor (b) và Virginia (c), mỗi loại chứa 50 mẫu. Mỗi mẫu chứa bốn thuộc tính, đó là chiều dài calax (SL), chiều rộng calax (SW), chiều dài cánh hoa (PL) và chiều rộng cánh hoa (PW). Một nhóm dữ liệu theo FOD không đầy đủ, được hiển thị trong Bảng 14, được thu thập sau quá trình tạo BPA trong [63].

Bảng 14. Kết quả mô hình hóa dữ liệu cho tròng đen có hàm khối lượng

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Attribute** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| SL | 0.6800 | 0.0000 | 0.0000 | 0.0610 | 0.0000 | 0.0000 | 0.1330 | 0.1260 |
| SW | 0.5030 | 0.0000 | 0.0000 | 0.0000 | 0.0100 | 0.0000 | 0.0000 | 0.4870 |
| PL | 0.9200 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0800 |
| PW | 0.8650 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.1350 |

Sử dụng phương pháp được đề xuất để tính toán các giá trị của EBBF, chúng ta có thể nhận được:

neb(Ri)SL = 0.1457, neb(Ri)SW = 0.1569, neb(Ri)PL = 0.1446, neb(Ri)PW = 0.1459 (57)

Sau đó, GCR được sử dụng để hợp nhất để thu được kết quả. Đồng thời, các BPA từ các bằng chứng trong bảng được hợp nhất trực tiếp với GCR và kết quả của hai phương pháp được so sánh, như thể hiện trong Bảng 15.

Bảng 15. Kết quả của hai phương pháp kết hợp trong ứng dụng để phân loại.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Fusion Methods** | **m({a})** | **m({b})** | **m({c})** | **m({a,b})** | **m({a,c})** | **m({b,c})** | **m({a,b,c})** | **m({∅})** |
| The proposed method | 0.9207 | 0.0363 | 0.0339 | 0.0031 | 0.0027 | 0.0027 | 0.0002 | 0.0005 |

Các kết quả hợp nhất được hiển thị trong Bảng 15. Kết quả cho thấy phương pháp này có hiệu suất hợp nhất tốt. Trong FOD không đầy đủ, lớp a chính xác vẫn được đưa ra một giá trị niềm tin rất cao, đó là kết quả trực quan. Thí nghiệm cho thấy phương pháp này cũng hợp lý và hiệu quả trong thế giới thực.

1. **Kết luận**

Trong FOD toàn diện, khi DCR được sử dụng để hợp nhất dữ liệu có tính mâu thuẫn cao, kết quả thường trái ngược với trực giác. Để giải quyết vấn đề này, một số học giả đưa ra chức năng niềm tin cơ bản để sửa đổi BPA nhằm loại bỏ tính tuyệt đối do xung đột mang lại. Theo giả thuyết FOD không đầy đủ, quy tắc kết hợp tổng quát có các vấn đề tương tự như DCR, nhưng chức năng niềm tin cơ bản không thể được áp dụng cho FOD không đầy đủ. Trong bài báo này, một hàm niềm tin cơ bản mở rộng dựa trên hàm niềm tin cơ bản và một phương pháp BPA sửa đổi tương ứng được đề xuất. Phương pháp này không chỉ tương thích với chức năng niềm tin cơ bản và có thể được sử dụng để quản lý dữ liệu xung đột trong FOD toàn diện, mà còn có thể được áp dụng hiệu quả trong FOD không đầy đủ. Sau khi sửa đổi BPA bằng EBBF, tính tuyệt đối do dữ liệu xung đột cao mang lại có thể được loại bỏ một cách hiệu quả. Phương pháp này không chỉ xem xét tiềm năng của FOD, mà còn xem xét các nguồn thông tin không chắc chắn không được tính đến bởi hàm niềm tin cơ sở và các phương pháp hiện có khác, bao gồm thông tin không chắc chắn do hàm khối lượng nonzero của tập rỗng và FOD không hoàn chỉnh mang lại. Để xác minh việc áp dụng phương pháp được đề xuất trong thực tế, bài báo này cũng thiết kế một phương pháp quản lý dữ liệu xung đột dựa trên EBBF và thảo luận và xác minh tính khả thi và hiệu quả của phương pháp thông qua một số ví dụ và ứng dụng.

Tuy nhiên, phương pháp được đề xuất trong bài báo này vẫn còn một số vấn đề mở đáng để thảo luận. Đầu tiên là vấn đề phức tạp tính toán. Phương pháp được đề xuất trong bài báo này có số lượng tính toán lớn như GCR, nhưng nó có thể được áp dụng cho nhiều tình huống hơn GCR. Nói tóm lại, phương pháp được đề xuất trong bài báo này không đơn giản, nhưng nó có hiệu quả. Sau đó, vì phương pháp này được đề xuất trong điều kiện thông tin không đầy đủ và trong FOD không đầy đủ, có nhiều nguồn thông tin không chắc chắn, hàm khối lượng tập rỗng tượng trưng được sử dụng trong phương pháp này. Trong các cải tiến tiếp theo, các nguồn thông tin không chắc chắn khác có thể được xem xét.