ĐỘ ĐO TƯƠNG TỰ MỚI TRÊN CÁC TẬP MỜ BỨC TRANH VÀ ỨNG DỤNG TRONG PHÂN CỤM DỮ LIỆU

Lê Thị Diệu Thùy*, Nguyễn Hữu Hải, Nguyễn Văn Hạnh, Đỗ Thị Huệ

Khoa Công nghệ thông tin, Học viện Nông Nghiệp Việt Nam

Tác giả liên hệ: ltdthuy@vnua.edu.vn

Ngày nhận bài: 08.07.2019 Ngày chấp nhận đăng: 26.08.2019

TÓM TẮT

Chỉ số Jaccard là một chỉ số trong thống kê dùng để so sánh độ giống nhau và sự đa dạng giữa các bộ mẫu. Trong bài báo này chúng tôi đề xuất một độ đo tương tự mới giữa các tập mờ bức tranh dựa trên chỉ số Jaccard. Sau đó chúng tôi đưa ra một số ví dụ cho thấy độ đo tương tự mới đã khắc phục được những hạn chế của các độ đo tương tự đã có. Cuối cùng chúng tôi sử dụng độ đo tương tự mới vào bài toán phân cụm dữ liệu.

Từ khóa: Tập mờ bức tranh, độ đo tương tự, bài toán phân cụm.

A New Similarity Measure of Picture Fuzzy Sets and Its Application to Data Clustering

ABSTRACT

The Jaccard index is a statistic used for comparing the similarity and diversity of sample sets. In this paper, we proposed a new similarity measure for picture fuzzy sets based on the Jaccard index. We then compared the proposed similarity measure with some existing similarity measures and showed that the new similarity measure overcomes the restrictions of the existing similarity measures. Finally, we used this new similarity measure for the data clustering problem.

Keywords: Picture fuzzy set, similarity measure, fuzzy clustering.

1. ĐẶT VẤN ĐỀ

Zadel (1965) lần đầu tiên đưa ra khái niệm và lý thuyết về tập mờ thông qua bài báo “Fuzzy Set” được đăng trên tạp chí Information and Control, đã mở đầu cho sự phát triển và ứng dụng của lý thuyết này. Ngày nay lý thuyết tập mờ vẫn không ngừng phát triển và đã được ứng dụng trong nhiều lĩnh vực nghiên cứu như lý thuyết điều khiển, trí tuệ nhân tạo, khai phá dữ liệu,... Định nghĩa tập mờ của Zadel sử dụng một hàm thuộc để mô tả cho mức độ của một phần tử thuộc về một tập. Atanasov (1986) đã mở rộng khái niệm tập mờ bằng khái niệm tập mờ trực cảm (Intuitionistic fuzzy sets), ngoài hàm thuộc, ông sử dụng thêm một hàm không thuộc để biểu thị độ không thuộc của một phần tử vào tập hợp. Bùi Công Cường (2014) giới thiệu khái niệm tập mờ bức tranh với ba hàm thành viên là hàm thuộc khẳng định, hàm thuộc phủ định và hàm thuộc trung lập. Về cơ bản, lý thuyết mờ bức tranh phù hợp với các tình huống khi một vấn đề có nhiều câu trả lời, khi đó lý thuyết tập mờ và tập mờ trực cảm không giải quyết được. Chẳng hạn trong các tình huống tổng hợp ý kiến của mọi người về một vấn đề trong đó có 4 câu trả lời cơ bản: có, không, không biết và không đưa ra câu trả lời. Bầu cử là một ví dụ điển hình, người bỏ phiếu được phân làm bốn nhóm: ủng hộ, phản đối, bỏ phiếu trắng hoặc phiếu không hợp lệ và không bỏ phiếu. Hiện nay lý thuyết tập mờ bức tranh đã và đang được các nhà nghiên cứu tiếp tục tìm hiểu, khai thác và có nhiều ứng dụng trong thực tiễn. Phạm Huy Thông & Lê Hoàng Sơn (2014) đã phát triển mô hình lai mới giữa tập mờ bức tranh và tập mờ trực cảm để chuẩn đoán y tế và ứng dụng vào hệ thống chăm sóc sức khỏe. Nguyễn Đình Hóa & cs. (2014) đề xuất phương pháp mới để dự báo thời tiết từ hình ảnh vệ tinh bằng cách sử dụng kết hợp phân cụm mờ bức tranh và hồi quy. Phạm Hồng Phong & cs. (2014) đã kiểm tra một số tính chất của phép hợp thành quan hệ mờ bức tranh và đề xuất một cách tiếp cận mới trong chuẩn đoán y khoa bằng cách sử dụng phép hợp thành quan hệ mờ bức tranh. Bùi Công Cường và Phạm Văn Hải (2015) nghiên cứu các toán tử logic mờ. Phạm Văn Việt & cs. (2015) đã đưa ra hệ thống suy luận mờ bức tranh dựa vào biểu đồ thành viên. Singh (2015) đề xuất hệ số tương quan cho tập mờ bức tranh và ứng dụng hệ số tương quan vào bài toán phân cụm mờ bức tranh. Lê Hoàng Sơn (2015) giới thiệu một số thuật toán phân cụm mờ bức tranh và ứng dụng trong dự báo thời tiết, dự báo chuỗi thời gian. Nguyễn Xuân Thảo & Nguyễn Văn Định (2015) đưa ra khái niệm tập mờ bức tranh - thô và nghiên cứu cấu trúc tô pô của tập mờ bức tranh - thô. Nguyễn Văn Định & cs. (2015) nghiên cứu về cơ sở dữ liệu mờ bức tranh. Lê Hoàng Sơn (2016) đưa ra độ đo khoảng cách tổng quát cho các tập mờ bức tranh và ứng dụng trong phân cụm. Nguyễn Đình Hóa & cs. (2017) đưa ra một số cải tiến cho thuật toán phân cụm mờ bằng cách sử dụng các tập mờ bức tranh và ứng dụng trong phân cụm dữ liệu địa lý. Phạm huy Thông (2016; 2017) đã có nhiều nghiên cứu về phân cụm dữ liệu mờ bức tranh. Garg (2017) trình bày một số toán tử tổng hợp mờ bức tranh và ứng dụng trong quyết định đa tiêu chí. Lê Hoàng Sơn & cs. (2017) đã đề xuất hệ thống suy luận mới trên tập mờ bức tranh. Lê Hoàng Sơn (2017) đưa các độ đo mới trên các tập mờ bức tranh, đó là các độ đo khoảng cách tổng quát và các độ đo kết hợp bức tranh. Peng & Dai (2017) đề xuất độ đo khoảng cách mới cho các tập mờ bức tranh và đưa ra thuật toán cho bài toán quyết định đa tiêu chí mờ bức tranh dựa trên các độ đo này. Nguyễn Văn Định & Nguyễn Xuân Thảo (2018) đề xuất các độ đo khoảng cách, độ đo không tương tự trên tập mờ bức tranh và ứng dụng trong quyết định đa tiêu chí. Phạm Thị Minh Phương & cs. (2018) nghiên cứu cơ sở lý thuyết cho thuật toán phân cụm mờ bức tranh. Zeng & cs. (2019) đề xuất độ đo phân kỳ mũ Jensen cho các tập mờ bức tranh và ứng dụng trong quyết định đa tiêu chí…

Độ đo tương tự là một công cụ hữu ích để xác định sự giống nhau giữa hai đối tượng. Nhiều độ đo tương tự khác nhau trên các tập mờ trực cảm đã được nghiên cứu, như các nghiên cứu của Szmidt & Kacorhot (2000), Li & Cheng (2002), Dengfeng (2002), Mitchell (2003), Liu (2005), Szmidt (2005), Hung & Yang (2007), Xu & Xia (2010), Ye (2011, 2012), Shi & Ye (2013), Tian (2013), Szmidt (2014), Ye (2016), Sơn & Phong (2016), Wei (2017),… Trên các tập mờ bức tranh cũng đã có một số độ đo tương tự được đề xuất và ứng dụng trong các bài toán khác nhau. Wei (2017, 2018) đề xuất một số độ đo tương tự cho tập mờ bức tranh và ứng dụng trong bài toán ra quyết định. Joshi & Kumar (2018) đưa ra độ đo tương tự cho tập mờ bức tranh dựa trên chỉ số Dice. Wei (2018) đưa ra một số độ đo tương tự Dice tổng quát… Tuy nhiên các độ đo này vẫn còn một số hạn chế, những hạn chế đó sẽ được chỉ ra ở phần 3 của bài báo. Chỉ số Jaccard là chỉ số thống kê dùng để so sánh sự giống nhau và sự đa dạng của các bộ mẫu. Dựa vào chỉ số này gần đây, Hwang & cs. (2018) đã đưa ra một số độ đo tương tự mới trên tập mờ trực cảm. Trên cơ sở đó, chúng tôi muốn nghiên cứu đề xuất độ đo tương tự mới trên tập mờ bức tranh dựa trên chỉ số Jaccard và ứng dụng độ đo này trong bài toán phân cụm.

2. PHƯƠNG PHÁP NGHIÊN CỨU

Trong bài báo này chúng tôi sử dụng phương pháp nghiên cứu lý thuyết, phương pháp phân tích và tổng hợp lý thuyết thông qua các tài liệu về tập mờ, tập mờ bức tranh, độ đo tương tự trên tập mờ bức tranh, các thuật toán phân cụm dữ liệu. Chúng tôi nghiên cứu các độ đo tương tự đã có trên tập mờ bức tranh, phân tích ưu nhược điểm của chúng, từ đó đề xuất độ đo tương tự mới trên tập mờ bức tranh dựa trên chỉ số Jaccard và tính toán với các ví dụ để so sánh độ đo mới với các độ đo đã có. Sau đó chúng tôi xây dựng thuật toán phân cụm cho tập mờ bức tranh dựa ma trận kết hợp tương đương và áp dụng độ đo mới vào thuật toán này.

Dưới đây là các khái niệm cơ bản về tập mờ bức tranh, độ đo tương tự trên tập mờ bức tranh và chỉ số Jaccard.

2.1. Tập mờ bức tranh

Bùi Công Cường (2014) đã đưa ra khái niệm tập mờ bức tranh như sau:

Định nghĩa 1. Cho tập nền $X=\{x_1,x_2,...,x_n\}$ một tập mờ bức tranh A trên X được xác định bởi

$$A=\left\{\left\langle x,\mu_A(x),\eta_A(x),\nu_A(x)\right\rangle \middle| x\in X\right\}$$

Với μ_A: X → [0;1] là hàm thuộc khẳng định, η_A: X → [0;1] là hàm thuộc trung lập, ν_A: X → [0;1] là hàm thuộc phủ định và thỏa mãn điều kiện $\mu_A(x) + \eta_A(x) + \nu_A(x) \le 1\ \forall x \in X$

Ta kí hiệu PFS(X) là tập hợp tất cả các tập mờ bức tranh trên X.

Định nghĩa 2. Cho hai tập mờ bức tranh $A,B \in PFS(X)$:

(1) $A\subseteq B \Leftrightarrow \mu_A(x)\le\mu_B(x),\ \eta_A(x)\le\eta_B(x),\ \nu_A(x)\ge\nu_B(x)\ \forall x\in X$

(2) $A=B \Leftrightarrow \mu_A(x)=\mu_B(x),\ \eta_A(x)=\eta_B(x),\nu_A(x)=\nu_B(x)\ \forall x\in X$

2.2. Độ đo tương tự trên tập mờ bức tranh

Định nghĩa 3. Cho hai tập mờ bức tranh $A,B \in PFS(X)$. Ánh xạ $S: PFS(X)\times PFS(X)\to[0;1]$ gọi là độ đo tương tự giữa A, B nếu S(A, B) thỏa mãn các điều kiện sau:

(1) $0\le S(A,B)\le 1$

(2) $S(A,A)=1$

(3) $S(A,B)=S(B,A)$

(4) Nếu $A\subseteq B\subseteq C$ thì $S(A,C)\le S(A,B)$; $S(A,C)\le S(B,C)\ \forall A,B,C\in PFS(X)$

Một số độ đo tương tự trên tập mờ bức tranh đã được đề xuất:

Độ đo tương tự Cosine (Wei, 2017) theo công thức (1).

Độ đo tương tự dựa trên lý thuyết tập hợp (Wei, 2018) theo công thức (2)

Độ đo tương tự Grey (Wei, 2018) theo công thức (3).

Các độ đo tương tự dựa trên hàm cosin (Wei, 2017) theo công thức (4) và (5).

Độ đo tương tự dựa trên hàm cotan (Wei, 2017) theo công thức (6).

Độ đo tương tự Dice (Joshi & Kumar, 2018) theo công thức (7)

$$PFC^1(A,B)=\frac{1}{n}\sum_{i=1}^{n}\frac{\mu_A(x_i)\mu_B(x_i)+\eta_A(x_i)\eta_B(x_i)+\nu_A(x_i)\nu_B(x_i)}{\sqrt{\mu_A^2(x_i)+\eta_A^2(x_i)+\nu_A^2(x_i)}\sqrt{\mu_B^2(x_i)+\eta_B^2(x_i)+\nu_B^2(x_i)}} \quad (1)$$

$$PFC^2(A,B)=\frac{1}{n}\sum_{i=1}^{n}\frac{\mu_A(x_i)\mu_B(x_i)+\eta_A(x_i)\eta_B(x_i)+\nu_A(x_i)\nu_B(x_i)}{\max(\mu_A^2(x_i)+\eta_A^2(x_i)+\nu_A^2(x_i),\mu_B^2(x_i)+\eta_B^2(x_i)+\nu_B^2(x_i))} \quad (2)$$

$$PFC^3(A,B)=\frac{1}{3n}\sum_{i=1}^{n}\left(\frac{\Delta\mu_{min}+\Delta\mu_{max}}{\Delta\mu_i+\Delta\mu_{max}}+\frac{\Delta\eta_{min}+\Delta\eta_{max}}{\Delta\eta_i+\Delta\eta_{max}}+\frac{\Delta\nu_{min}+\Delta\nu_{max}}{\Delta\nu_i+\Delta\nu_{max}}\right) \quad (3)$$

Với: $\Delta\mu_i=|\mu_A(x_i)-\mu_B(x_i)|$, $\Delta\mu_{min}=\min_i\{|\mu_A(x_i)-\mu_B(x_i)|\}$, $\Delta\mu_{max}=\max_i\{|\mu_A(x_i)-\mu_B(x_i)|\}$

$\Delta\eta_i=|\eta_A(x_i)-\eta_B(x_i)|$, $\Delta\eta_{min}=\min_i\{|\eta_A(x_i)-\eta_B(x_i)|\}$, $\Delta\eta_{max}=\max_i\{|\eta_A(x_i)-\eta_B(x_i)|\}$

$\Delta\nu_i=|\nu_A(x_i)-\nu_B(x_i)|$, $\Delta\nu_{min}=\min_i\{|\nu_A(x_i)-\nu_B(x_i)|\}$, $\Delta\nu_{max}=\max_i\{|\nu_A(x_i)-\nu_B(x_i)|\}$.

$$PFCS^1(A,B)=\frac{1}{n}\sum_{i=1}^{n}\cos\left\{\frac{\pi}{2}[\,|\mu_A(x_i)-\mu_B(x_i)|\vee|\eta_A(x_i)-\eta_B(x_i)|\vee|\nu_A(x_i)-\nu_B(x_i)|\,]\right\} \quad (4)$$

$$PFCS^2(A,B)=\frac{1}{n}\sum_{i=1}^{n}\cos\left\{\frac{\pi}{4}[\,|\mu_A(x_i)-\mu_B(x_i)|+|\eta_A(x_i)-\eta_B(x_i)|+|\nu_A(x_i)-\nu_B(x_i)|\,]\right\} \quad (5)$$

$$\mathrm{PFCT}(A,B)=\frac{1}{n}\sum_{i=1}^{n}\cot\left\{\frac{\pi}{4}+\frac{\pi}{4}[\,|\mu_A(x_i)-\mu_B(x_i)|+|\eta_A(x_i)-\eta_B(x_i)|+|\nu_A(x_i)-\nu_B(x_i)|\,]\right\} \quad (6)$$

$$D_{PFS}(A,B)=\frac{1}{n}\sum_{i=1}^{n}\frac{2\left(\mu_A(x_i)\mu_B(x_i)+\eta_A(x_i)\eta_B(x_i)+\nu_A(x_i)\nu_B(x_i)\right)}{\mu_A^2(x_i)+\eta_A^2(x_i)+\nu_A^2(x_i)+\mu_B^2(x_i)+\eta_B^2(x_i)+\nu_B^2(x_i)} \quad (7)$$

$$\begin{aligned}J_{PFS}(A,B)=\frac{1}{4n}\sum_{i=1}^{n}\Bigg(&\frac{e^{-\mu_A(x_i)}e^{-\mu_B(x_i)}}{e^{-2\mu_A(x_i)}+e^{-2\mu_B(x_i)}-e^{-\mu_A(x_i)}e^{-\mu_B(x_i)}}\\&+\frac{e^{-\eta_A(x_i)}e^{-\eta_B(x_i)}}{e^{-2\eta_A(x_i)}+e^{-2\eta_B(x_i)}-e^{-\eta_A(x_i)}e^{-\eta_B(x_i)}}\\&+\frac{e^{-(1-\nu_A(x_i))}e^{-(1-\nu_B(x_i))}}{e^{-2(1-\nu_A(x_i))}+e^{-2(1-\nu_B(x_i))}-e^{-(1-\nu_A(x_i))}e^{-(1-\nu_B(x_i))}}\\&+\frac{e^{-\frac{1}{2}(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))}e^{-\frac{1}{2}(1+\mu_B(x_i)+\eta_B(x_i)-\nu_B(x_i))}}{e^{-(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))}+e^{-(1+\mu_B(x_i)+\eta_B(x_i)-\nu_B(x_i))}-e^{-\frac{1}{2}(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))}e^{-\frac{1}{2}(1+\mu_B(x_i)+\eta_B(x_i)-\nu_B(x_i))}}\Bigg)\end{aligned} \quad (8)$$

2.3. Chỉ số Jaccard

Chỉ số Jaccard là chỉ số được sử dụng trong thống kê để đo mức độ tương tự giữa hai bộ mẫu. Cho hai bộ mẫu A và B, chỉ số Jaccard giữa A và B được xác định bởi công thức:

$$J(A,B)=\frac{|A\cap B|}{|A\cup B|}=\frac{|A\cap B|}{|A|+|B|-|A\cap B|}$$

Ta có $0\le J(A,B)\le 1$ và $J(A,B)$ càng lớn thì mức độ tương đồng giữa A, B càng lớn.

Nếu xét hai véc tơ n chiều là X = (x_1, x_2,..., x_n), Y = (y_1, y_2,..., y_n) thì chỉ số Jaccard giữa X_A, X_B xác định như sau:

$$J(X,Y)=\frac{XY}{\|X\|^2+\|Y\|^2-XY}=\frac{\sum_{i=1}^{n}x_iy_i}{\sum_{i=1}^{n}x_i^2+\sum_{i=1}^{n}y_i^2-\sum_{i=1}^{n}x_iy_i}$$

3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Độ đo tương tự mới dựa trên chỉ số Jaccard giữa các tập mờ bức tranh

Dựa trên chỉ số Jaccard, chúng tôi đề xuất độ đo tương tự giữa các tập mờ bức tranh bằng cách đo mức tương tự giữa ba hàm thành viên trên hai tập mờ bức tranh như sau:

Cho:

$A=\left\{\left\langle x_j,(\mu_A(x_j),\eta_A(x_j),\nu_A(x_j))\right\rangle \mid x_j\in X\right\}$, $B=\left\{\left\langle x_j,(\mu_B(x_j),\eta_B(x_j),\nu_B(x_j))\right\rangle \mid x_j\in X\right\}$ là hai tập mờ bức tranh trên tập nền $X=\{x_1,x_2,...,x_n\}$. Khi đó độ đo tương tự giữa A, B được xác định như sau như công thức (8).

Ta chứng minh $J_{PFS}(A,B)$ thỏa mãn 4 tính chất của độ đo tương tự:

Xét hàm $f(x,y)=\dfrac{xy}{x^2+y^2-xy}$. Ta có :

$$\begin{aligned}J_{PFS}(A,B)=\frac{1}{4n}\sum_{i=1}^{n}\Big[&f\left(e^{-\mu_A(x_i)},e^{-\mu_B(x_i)}\right)\\&+f\left(e^{-\eta_A(x_i)},e^{-\eta_B(x_i)}\right)\\&+f\left(e^{-(1-\nu_A(x_i))},e^{-(1-\nu_B(x_i))}\right)\\&+f\left(e^{-\frac{1}{2}(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))},e^{-\frac{1}{2}(1+\mu_B(x_i)+\eta_B(x_i)-\nu_B(x_i))}\right)\Big]\end{aligned}$$

(1) Dễ thấy $0\le f(x,y)\le 1\ \forall x,y\ge 0$, do đó

$0\le f\left(e^{-\mu_A(x_i)},e^{-\mu_B(x_i)}\right)\le 1,$

$0\le f\left(e^{-\eta_A(x_i)},e^{-\eta_B(x_i)}\right)\le 1,$

$0\le f\left(e^{-(1-\nu_A(x_i))},e^{-(1-\nu_B(x_i))}\right)\le 1,$

$0\le f\left(e^{-\frac{1}{2}(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))},e^{-\frac{1}{2}(1+\mu_B(x_i)+\eta_B(x_i)-\nu_B(x_i))}\right)\le 1\ \forall i=1,2,...,n$

Vậy $0 \le J_{PFS}(A,B) \le 1$.

(2) Giả sử A = B, tức là $\mu_A(x_i) = \mu_B(x_i)$, $\eta_A(x_i) = \eta_A(x_i)$, $\nu_A(x_i) = \nu_B(x_i)$.

Khi đó

$$e^{-\mu_A(x_i)} = e^{-\mu_B(x_i)}, e^{-\eta_A(x_i)} = e^{-\eta_B(x_i)}, e^{-(1-\nu_A(x_i))} = e^{-(1-\nu_B(x_i))}, e^{-\frac{1}{2}(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))} = e^{-\frac{1}{2}(1+\mu_B(x_i)+\eta_B(x_i)-\nu_B(x_i))}$$

Mặt khác nếu x = y thì f(x,y) =1. Do đó:

$$f\left(e^{-\mu_A(x_i)}, e^{-\mu_B(x_i)}\right) = 1,$$

$$f\left(e^{-\eta_A(x_i)}, e^{-\eta_B(x_i)}\right) = 1,$$

$$f\left(e^{-(1-\nu_A(x_i))}, e^{-(1-\nu_B(x_i))}\right) = 1,$$

$$f\left(e^{-\frac{1}{2}(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))}, e^{-\frac{1}{2}(1+\mu_B(x_i)+\eta_B(x_i)-\nu_B(x_i))}\right) = 1 \quad \forall i = \overline{1,n}$$

Suy ra $J_{PFS}(A,B) = 1$.

Ngược lại giả sử $J_{PFS}(A,B) = 1$, khi đó

$$f\left(e^{-\mu_A(x_i)}, e^{-\mu_B(x_i)}\right) = 1,$$

$$f\left(e^{-\eta_A(x_i)}, e^{-\eta_B(x_i)}\right) = 1,$$

$$f\left(e^{-(1-\nu_A(x_i))}, e^{-(1-\nu_B(x_i))}\right) = 1, ,$$

$$f\left(e^{-\frac{1}{2}(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))}, e^{-\frac{1}{2}(1+\mu_B(x_i)+\eta_B(x_i)-\nu_B(x_i))}\right) = 1 \ \forall i = \overline{1,n}$$

Mà $f(x,y) = 1 \Leftrightarrow \dfrac{xy}{x^2+y^2-xy} = 1 \Leftrightarrow (x-y)^2 = 0 \Leftrightarrow x = y$

Từ đó ta có $\mu_A(x_i) = \mu_B(x_i), \eta_A(x_i) = \eta_B(x_i)$, $\nu_A(x_i) = \nu_B(x_i)$, hay A= B.

(3) Dễ thấy $J_{PFS}(A,B) = J_{PFS}(B,A)$.

(4) Ta có $f'(x) = \dfrac{y(y^2-x^2)}{(x^2+y^2-xy)^2}$. Do đó f(x) đồng biến trên (0,y) và nghịch biến trên (y,1).

Giả sử $A \subseteq B \subseteq C$, tức là với $\forall i = \overline{1,n}$

$$\mu_A(x_i) \le \mu_B(x_i) \le \mu_C(x_i),\ \eta_A(x_i) \le \eta_B(x_i) \le \eta_C(x_i),\ \nu_A(x_i) \ge \nu_B(x_i) \ge \nu_C(x_i)$$

Suy ra $e^{-\mu_A(x_i)} \ge e^{-\mu_B(x_i)} \ge e^{-\mu_C(x_i)}$, $e^{-\eta_A(x_i)} \ge e^{-\eta_B(x_i)} \ge e^{-\eta_C(x_i)}$, $e^{-(1-\nu_A(x_i))} \ge e^{-(1-\nu_B(x_i))} \ge e^{-(1-\nu_C(x_i))}$ và $e^{-\frac{1}{2}(\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))} \ge e^{-\frac{1}{2}(\mu_B(x_i)+\eta_B(x_i)-\nu_C(x_i))} \ge e^{-\frac{1}{2}(\mu_C(x_i)+\eta_C(x_i)-\nu_C(x_i))}$.

Do đó

$$f\left(e^{-\mu_A(x_i)}, e^{-\mu_C(x_i)}\right) \le f\left(e^{-\mu_A(x_i)}, e^{-\mu_B(x_i)}\right),$$

$$f\left(e^{-\eta_A(x_i)}, e^{-\eta_C(x_i)}\right) \le f\left(e^{-\eta_A(x_i)}, e^{-\eta_B(x_i)}\right),$$

$$f\left(e^{-(1-\nu_A(x_i))}, e^{-(1-\nu_C(x_i))}\right) \le f\left(e^{-(1-\nu_A(x_i))}, e^{-(1-\nu_B(x_i))}\right),$$

$$f\left(e^{-\frac{1}{2}(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))}, e^{-\frac{1}{2}(1+\mu_C(x_i)+\eta_C(x_i)-\nu_C(x_i))}\right) \le f\left(e^{-\frac{1}{2}(1+\mu_A(x_i)+\eta_A(x_i)-\nu_A(x_i))}, e^{-\frac{1}{2}(1+\mu_B(x_i)+\eta_B(x_i)-\nu_B(x_i))}\right)$$

Do đó ta có $J_{PFS}(A,C) \le J_{PFS}(A,B)$. Tương tự có $J_{PFS}(A,C) \le J_{PFS}(B,C)$.

3.2. Một số ví dụ

Dưới đây là một số ví dụ so sánh độ đo tương tự mới với các độ đo tương tự đã có.

Ví dụ 1. Xét bài toán y tế trong Dutta (2017) như sau: giả sử trên tập nền $X = \{x_1, x_2, x_3, x_4, x_5\}$ có 5 tập mờ bức tranh

$$A_1 = \{\langle x_1, 0.4, 0, 0\rangle, \langle x_2, 0.3, 0.2, 0.4\rangle, \langle x_3, 0.1, 0.35, 0.5\rangle, \langle x_4, 0.4, 0.3, 0.2\rangle, \langle x_5, 0.1, 0.25, 0.5\rangle\}$$

$$A_2 = \{\langle x_1, 0.7, 0, 0\rangle, \langle x_2, 0.2, 0.4, 0.35\rangle, \langle x_3, 0, 0.4, 0.5\rangle, \langle x_4, 0.7, 0.1, 0\rangle, \langle x_5, 0.1, 0.3, 0.5\rangle\}$$

$$A_3 = \{\langle x_1, 0.3, 0.4, 0.3\rangle, \langle x_2, 0.6, 0.2, 0.1\rangle, \langle x_3, 0.2, 0.3, 0.4\rangle, \langle x_4, 0.2, 0.35, 0.3\rangle, \langle x_5, 0.1, 0.2, 0.6\rangle\},$$

$$A_4 = \{\langle x_1, 0.1, 0.3, 0.5\rangle, \langle x_2, 0.2, 0.4, 0.3\rangle, \langle x_3, 0.8, 0, 0\rangle, \langle x_4, 0.2, 0.4, 0.3\rangle, \langle x_5, 0.2, 0.35, 0.3\rangle\}$$

$$A_5 = \{\langle x_1, 0.1, 0.3, 0.5\rangle, \langle x_2, 0, 0.5, 0.35\rangle \langle x_3, 0.2, 0.3, 0.5\rangle, \langle x_4, 0.2, 0.35, 0.4\rangle, \langle x_5, 0.8, 0, 0.1\rangle\}$$

Giả sử có tập mờ bức tranh mới là:

$$B = \{\langle x_1, 0.8, 0, 0.1\rangle, \langle x_2, 0.6, 0.3, 0.1\rangle, \langle x_3 = 0.2, 0.4, 0.4\rangle, \langle x_4 = 0.6, 0.15, 0.1\rangle, \langle x_3 = 0.1, 0.4, 0.4\rangle\}$$

Sử dụng công thức (1), (2), (3), (4), (5), (6), (7), (8) để đo độ tương tự giữa B và A_i (i =1, 2, 3, 4, 5) ta được kết quả như trong bảng 1.

Như vậy độ đo mới có cùng kết quả với tất cả các độ đo còn lại, đó là độ tương tự giữa B và A_2 là lớn nhất, do đó xếp B vào lớp của A_2.

Ví dụ 2. Xét dữ liệu trong Singh (2015) như sau: giả sử trên tập nền $X = \{x_1, x_2, x_3, x_4, x_5\}$ có 3 tập mờ bức tranh

$$A_1 = \{\langle x_1, 0.4, 0.5, 0.1\rangle, \langle x_2, 0.7, 0.1, 0.1\rangle, \langle x_3, 0.3, 0.3, 0.2\rangle\}$$

$$A_2 = \{\langle x_1, 0.5, 0.4, 0.1\rangle, \langle x_2, 0.7, 0.2, 0.1\rangle, \langle x_3, 0.4, 0.3, 0.1\rangle\},$$

$$A_3 = \{\langle x_1, 0.4, 0.5, 0.1\rangle, \langle x_2, 0.7, 0.1, 0.1\rangle, \langle x_3, 0.4, 0.3, 0.2\rangle\}$$

Giả sử có tập mờ bức tranh mới là:

$$B = \{\langle x_1, 0.1, 0.1, 0.4\rangle, \langle x_2, 1, 0, 0\rangle, \langle x_3 = 0, 1, 0\rangle\}$$

Sử dụng công thức (1), (2), (3), (4), (5), (6), (7), (8) để đo độ tương tự giữa A_1 và B, giữa A_2 và B, giữa A_3 và B ta được kết quả như trong bảng 2.

Bảng 1. So sánh độ đo tương tự và xếp hạng kết quả phân lớp trong ví dụ 1

Độ đo tương tự	Độ đo tương tự giữa B và A_i = (i = 1, 2)					Kết luận
	A_1	A_2	A_3	A_4	A_4	
PFC^1	0.9167	0.9193	0.8191	0.5875	0.5098	B vào lớp của A_2
PFC^2	0.7517	0.8301	0.712	0.4841	0.4551	B vào lớp của A_2
PFC^3	0.7703	0.8089	0.7899	0.7571	0.7336	B vào lớp của A_2
$PFCS^1$	0.9222	0.9446	0.8885	0.7295	0.6585	B vào lớp của A_2
$PFCS^2$	0.9353	0.952	0.8807	0.7053	0.667	B vào lớp của A_2
PFCT	0.6983	0.7597	0.6717	0.4876	0.4337	B vào lớp của A_2
D_{PFS}	0.8701	0.9132	0.8077	0.5754	0.5033	B vào lớp của A_2
J_{PFS}	0.969	0.9786	0.96	0.9063	0.8921	B vào lớp của A_2

Bảng 2. So sánh độ đo tương tự và xếp hạng kết quả phân lớp trong ví dụ 2

Độ đo tương tự	Độ đo tương tự giữa B và A_i (i =1, 2)			Kết luận
	A_1	A_2	A_3	
PFC^1	0.6975	0.6712	0.67	B vào lớp của A_1
PFC^2	0.4365	0.4365	0.4365	Null
PFC^3	0.8628	0.8844	0.8489	B vào lớp của A_2
$PFCS^1$	0.718	0.718	0.718	Null
$PFCS^2$	0.7396	0.7286	0.7178	B vào lớp của A_1
PFCT	0.4541	0.4541	0.4541	Null
D_{PFS}	0.6174	0.6062	0.6085	B vào lớp của A_1
J_{PFS}	0.9019	0.9001	0.8892	B vào lớp của A_1

Ghi chú: Null nghĩa là ta không có kết quả xếp lớp.

Như vậy các độ đo PFC^2, $PFCS^1$, PFCT không có kết quả phân lớp, độ đo PFC^3 cho kết quả B vào lớp của A_2, còn độ đo mới có cùng kết quả với các độ đo PFC^1, $PFCS^2$, D_{PFS}, đó là B vào lớp của A_1.

Ví dụ 3. Giả sử trên tập nền $X = \{x_1, x_2, x_3\}$ có 2 tập mờ bức tranh:

$$A_1 = \{\langle x_1, 0.3, 0.5, 0.2\rangle, \langle x_2, 0.1, 0.1, 0.5\rangle, \langle x_3, 0.6, 0.1, 0.1\rangle\},$$

$$A_2 = \{\langle x_1, 0.15, 0.3, 0.45\rangle, \langle x_2, 0.3, 0.3, 0.3\rangle, \langle x_3, 0.6, 0.15, 0.15\rangle\}$$

Giả sử có một tập mờ bức tranh mới là:

$$B = \{\langle x_1, 0.4, 0.4, 0.1\rangle, \langle x_2, 0.2, 0.2, 0.4\rangle, \langle x_3, 0.35, 0.2, 0.35\rangle\}.$$

Sử dụng công thức (1), (2), (3), (4), (5), (6), (7), (8) để đo độ tương tự giữa A_1 và B, giữa A_2 và B ta được kết quả như trong bảng 3.

Như vậy chỉ có độ đo mới J_{PFS} phân lớp được B vào nhóm nào, cụ thể B vào lớp của A_1.

Ví dụ 4. Giả sử trên tập nền $X = \{x_1, x_2, x_3\}$ có 2 tập mờ bức tranh:

$$A_1 = \{\langle x_1, 0.4, 0.4, 0.1\rangle, \langle x_2, 0.3, 0.4, 0.1\rangle, \langle x_3, 0.1, 0.3, 0.4\rangle\},$$

$$A_2 = \{\langle x_1, 0.5, 0.2, 0.2\rangle, \langle x_2, 0.5, 0.1, 0.0\rangle, \langle x_3, 0.4, 0.1, 0.3\rangle\}$$

Giả sử có một tập mờ bức tranh mới là:

$$B = \{\langle x_1, 0.3, 0.3, 0.3\rangle, \langle x_2, 0.3, 0.1, 0.3\rangle, \langle x_3, 0.2, 0.2, 0.2\rangle\}$$

Sử dụng công thức (1), (2), (3), (4), (5), (6), (7), (8) để đo độ tương tự giữa A_1 và B, giữa A_2 và B ta được kết quả như trong bảng 4.

Như vậy chỉ có độ đo mới phân lớp được B vào nhóm nào, cụ thể B thuộc vào nhóm của A_2.

Bảng 3. So sánh độ đo tương tự và xếp hạng kết quả phân lớp trong ví dụ 3

Độ đo tương tự	Độ đo tương tự giữa B và A_i (i=1, 2)		Kết luận
	A_1	A_2	
PFC^1	0.8209	0.8209	Null
PFC^2	0.5833	0.5833	Null
PFC^3	0.8329	0.8329	Null
$PFCS^1$	0.891	0.891	Null
$PFCS^2$	0.902	0.902	Null
PFCT	0.6128	0.6128	Null
D_{PFS}	0.7396	0.7396	Null
J_{PFS}	0.9672	0.9603	B vào lớp của A_1

Ghi chú: Null nghĩa là ta không có kết quả xếp lớp.

Bảng 4. So sánh độ đo tương tự và xếp hạng kết quả phân lớp trong ví dụ 4

Độ đo tương tự	Độ đo tương tự giữa B và A_i (i=1, 2)		Kết luận
	A_1	A_2	
PFC^1	0.8434	0.8434	Null
PFC^2	0.683	0.683	Null
PFC^3	0.8519	0.8519	Null
$PFCS^1$	0.931	0.931	Null
$PFCS^2$	0.942	0.942	Null
PFCT	0.6887	0.6887	Null
D_{PFS}	0.8177	0.8177	Null
J_{PFS}	0.9712	0.9744	B vào lớp của A_2

Ghi chú: Null nghĩa là ta không có kết quả xếp lớp.

3.3 Ứng dụng độ đo mới vào bài toán phân cụm dữ liệu

Xu & cs. (2018) đã đưa ra thuật toán phân cụm mờ trực cảm bằng ma trận kết hợp tương đương, trên cơ sở đó chúng tôi áp dụng độ đo tương tự mới vào bài toán phân cụm cho tập mờ bức tranh.

Định nghĩa 4. Cho $A_1, A_2, ..., A_n$ là các tập mờ bức tranh, khi đó $C = (c_{ij})_{n\times n}$ gọi là ma trận kết hợp mờ bức tranh nếu $c_{ij} = c(A_i, A_j)$ thỏa mãn các tính chất sau:

i) $0 \le c_{ij} \le 1 \ \forall i, j = 1,...,n$

ii) $c_{ij} = 1 \Leftrightarrow A_i = A_j$

iii) $c_{ij} = c_{ji} \ \forall i, j = 1,...,n.$

Định nghĩa 5. Cho $C = (c_{ij})_{n\times n}$ là một ma trận kết hợp. Khi đó ma trận $C^2 = C \circ C = \left(\overline{c_{ij}}\right)_{n\times n}$ gọi là ma trận hợp thành của C, với $\overline{c_{ij}} = \max_k \left\{\min\left\{c_{ik}, c_{kj}\right\}\right\}, i, j = 1,...,n$

Định lý 1. Cho $C = (c_{ij})_{n\times n}$ là một ma trận kết hợp. Khi đó ma trận hợp thành C^2 cũng là ma trận kết hợp.

Chứng minh:

i) Do $0 \le c_{ij} \le 1 \ \forall i, j = 1,...,n$ nên $0 \le \overline{c_{ij}} = \max_k \left\{\min\left\{c_{ik}, c_{kj}\right\}\right\} \le 1 \ \forall i, j = 1,...,n$

ii) Ta có $\overline{c_{ij}} = \max_k \left\{\min\left\{c_{ik}, c_{kj}\right\}\right\} = 1$ khi vào chỉ khi tồn tại $k \in \{1, 2, ..., n\}$ sao cho $c_{ik} = c_{kj}$. Do C là ma trận kết hợp nên điều này xảy ra khi và chỉ khi $A_i = A_k = A_j$.

iii) Do $c_{ij} = c_{ji} \ \forall i, j = 1,...,n$ nên:

$$\begin{aligned}\overline{c_{ij}} &= \max_k \left\{\min\left\{c_{ik}, c_{kj}\right\}\right\} \\ &= \max_k \left\{\min\left\{c_{ki}, c_{jk}\right\}\right\} \\ &= \max_k \left\{\min\left\{c_{jk}, c_{ki}\right\}\right\} = \overline{c_{ji}} \ \forall i, j = 1, 2, ..., n.\end{aligned}$$

Định lý 2. Cho $C = (c_{ij})_{n\times n}$ là một ma trận kết hợp. Khi đó với mọi số nguyên dương k, ma trận $C^{2^{k+1}}$ với $C^{2^{k+1}} = C^{2^k} \circ C^{2^k}$ cũng là ma trận kết hợp.

Chứng minh:

Với k = 0 ta có $C^2 = C \circ C$, do định lý 1 nên C^2 là ma trận kết hợp.

Giả sử $C^{2^{h+1}} = C^{2^h} \circ C^{2^h}$ là ma trận kết hợp. Ta chứng minh $C^{2^{h+2}}$ là ma trận kết hợp. Thật vậy theo Định nghĩa 5 ta có $C^{2^{h+2}} = C^{2^{h+1}} \circ C^{2^{h+1}}$, do đó theo Định lý 1 thì $C^{2^{h+2}}$ là ma trận kết hợp.

Định nghĩa 6. Cho $C = (c_{ij})_{n\times n}$ là một ma trận kết hợp. Nếu $C^2 \subseteq C$, tức là $\max_k \left\{\min\left\{c_{ik}, c_{kj}\right\}\right\} \le c_{ij} \ \forall i, j = 1,...,n$ thì C gọi là ma trận kết hợp tương đương.

Định lý 3. Cho $C = (c_{ij})_{n\times n}$ là một ma trận kết hợp. Khi đó dãy các ma trận kết hợp

$$C \to C^2 \to C^4 \to ... \to C^{2^k} \to ...$$

luôn tồn tại số nguyên dương k sao cho $C^{2^k} = C^{2^{k+1}}$, khi đó C^{2^k} là ma trận kết hợp tương đương.

Chứng minh

Với mỗi i, j xét dãy số $\left\{c_{ij}^{(2^k)}\right\}_{k=1}^{\infty}$, ở đó $c_{ij}^{(2^k)}$ là phần tử thuộc hàng i, cột j của ma trận C^{2^k}. Do cách xây dựng ma trận thành phần ở định nghĩa 5 ta có $\left\{c_{ij}^{(2^k)}\right\}_{k=1}^{\infty}$ là dãy không giảm và bị chặn, do đó tồn tại $c_{ij}^* = \lim_{k\to\infty} c_{ij}^{(2^k)}$.

Mặt khác dãy $\left\{c_{ij}^{(2^k)}\right\}_{k=1}^{\infty}$ chỉ nhận hữu hạn các giá trị khác nhau thuộc tập các phần tử của ma trận C, do đó với k đủ lớn thì $c_{ij}^{(2^k)} = c_{ij}^{(2^{k+1})}$.

Vậy với k đủ lớn thì $c_{ij}^{(2^k)} = c_{ij}^{(2^{k+1})} \ \forall i, j$, tức là $C^{2^k} = C^{2^{k+1}}$ và C^{2^k} là ma trận kết hợp tương đương.

Định nghĩa 7. Cho $C = (c_{ij})_{n\times n}$ là một ma trận kết hợp tương đương. Khi đó ma trận $C_\lambda = ({}_\lambda c_{ij})_{n\times n}$ được gọi là ma trận λ_lát cắt của C, ở đó

$$_\lambda c_{ij} = \begin{cases} 0, & c_{ij} < \lambda \\ 1, & c_{ij} \ge \lambda \end{cases} \quad i, j = 1, 2, ..., n.$$

λ gọi là mức tin cậy với $\lambda \in [0,1]$.

Sử dụng các khái niệm nêu trên chúng tôi đề xuất thuật toán phân cụm dựa trên ma trận kết hợp tương đương như sau:

Thuật toán phân cụm:

Đầu vào: Cho $A_1, A_2, ..., A_n$ là các tập mờ bức tranh trên tập nền $X = \{x_1, x_2, ..., x_n\}$

Đầu ra: kết quả phân tích cụm: là một phân hoạch của họ $\{A_1, A_2, ..., A_n\}$

Các bước của thuật toán:

Bước 1. Xây dựng ma trận kết hợp $C = (c_{ij})_{n\times n}$ với $c_{ij} = J_{PFS}(A_i, A_j)$ được tính với công thức (8).

Bước 2. Nếu $C = (c_{ij})_{n\times n}$ là ma trận kết hợp tương đương thì ta xây dựng ma trận λ_lát cắt $C_\lambda = ({}_\lambda c_{ij})_{n\times n}$ theo định nghĩa 7. Nếu không ta xây dựng dãy các ma trận $C \to C^2 \to C^4 \to ... \to C^{2^k} \to ...$ theo định lý 3 để có ma trận kết hợp tương đương $\overline{C}$, sau đó xây dựng ma trận lát cắt ${}_\lambda\overline{C}$ của $\overline{C}$.

Bước 3. Nếu mọi phần tử thuộc hàng (cột) thứ i của C_λ (hoặc ${}_\lambda\overline{C}$) giống các phần tử tương ứng thuộc hàng (cột) thứ j của C_λ (hoặc ${}_\lambda\overline{C}$) thì A_i, A_j được xếp vào một cụm.

Sau đây chúng ta xét ví dụ ứng dụng thuật toán phân cụm đã đề xuất ở trên.

Ví dụ 5. Xét bài toán phân cụm xe hơi (Singh, 2015), kí hiệu 4 chiếc xe hơi là A_1, A_2, A_3, A_4. Mỗi chiếc xe ta xét 3 thuộc tính: tiết kiệm nhiên liệu (x_1), giá cả (x_2), an toàn (x_3). Thông tin được thể hiện bằng tập mờ bức tranh như sau:

$$A_1 = \{\langle x_1, 0.3, 0.2, 0.1\rangle, \langle x_2, 0.8, 0.1, 0.1\rangle, \langle x_3, 0.1, 0, 0.9\rangle\}$$

$$A_2 = \{\langle x_1, 0.4, 0.1, 0.5\rangle, \langle x_2, 0, 0.7, 0.1\rangle, \langle x_3, 0.3, 0.2, 0.4\rangle\}$$

$$A_3 = \{\langle x_1, 0.2, 0.6, 0.1\rangle, \langle x_2, 0.9, 0.1, 0\rangle, \langle x_3, 0.7, 0.1, 0.2\rangle\}$$

$$A_4 = \{\langle x_1, 0.7, 0.3, 0\rangle, \langle x_2, 0.5, 0.1, 0.3\rangle, \langle x_3, 0.6, 0, 0.3\rangle\}$$

Bước 1. Sử dụng công thức (8) tính các $J_{PSF}(A_i, A_j)$, ta có ma trận kết hợp

$$C = \begin{pmatrix} 1 & 0.8582 & 0.9035 & 0.9055 \\ 0.8582 & 1 & 0.8750 & 0.8912 \\ 0.9035 & 0.8750 & 1 & 0.9414 \\ 0.9055 & 0.8912 & 0.9414 & 1 \end{pmatrix}$$

Bước 2. Tính ma trận hợp thành

$$C^2 = \begin{pmatrix} 1 & 0.8912 & 0.9055 & 0.9055 \\ 0.8912 & 1 & 0.8912 & 0.8912 \\ 0.9055 & 0.8912 & 1 & 0.9414 \\ 0.9055 & 0.8912 & 0.9414 & 1 \end{pmatrix}$$

Ta thấy C không phải là ma trận kết hợp tương đương

$$C^4 = \begin{pmatrix} 1 & 0.8912 & 0.9055 & 0.9055 \\ 0.8912 & 1 & 0.8912 & 0.8912 \\ 0.9055 & 0.8912 & 1 & 0.9414 \\ 0.9055 & 0.8912 & 0.9414 & 1 \end{pmatrix}$$

Ta có $C^4 = C^2$, do đó C^2 là ma trận kết hợp tương đương

Bước 3. Xây dựng ma trận λ_lát cắt của C^2 ta có các khả năng phân cụm như sau:

Với $\lambda \le 0.8912$: ta có một cụm $\{A_1, A_2, A_3, A_4\}$.

Với $0.8912 < \lambda < 0.905$: ta có hai cụm $\{A_1, A_3, A_4\}$, $\{A_2\}$.

Với $0.905 < \lambda \le 0.9414$: ta có ba cụm $\{A_1\}$, $\{A_2\}$, $\{A_3, A_4\}$.

Với $0.9414 < \lambda \le 1$: ta có 4 cụm $\{A_1\}$, $\{A_2\}$, $\{A_3\}$, $\{A_4\}$.

5. KẾT LUẬN

Trong bài báo này chúng tôi đã đề xuất một độ đo tương tự mới cho các tập mờ bức tranh dựa trên chỉ số Jaccard. Bằng một số ví dụ, chúng tôi đã chỉ ra rằng trong nhiều trường hợp việc sử dụng độ đo tương tự mới đã khắc phục được những hạn chế của các độ đo tương tự đã có. Chúng tôi cũng đã sử dụng độ đo tương tự mới vào bài toán phân cụm tập mờ bức tranh dựa trên thuật toán phân cụm ma trận kết hợp tương đương, điều này cho thấy tính khả dụng của độ đo tương tự mới được đề xuất. Trong tương lai độ đo tương tự được đề xuất cần được tiếp tục nghiên cứu ứng dụng nhiều hơn trong các bài toán nhận dạng mẫu, phân cụm, quyết định đa tiêu chí,… và nghiên cứu mở rộng hướng này trên các dạng tập mờ khác.

TÀI LIỆU THAM KHẢO

Atanassov K.T. (1986). Intuitionistic fuzzy sets. Fuzzy sets and Systems. 20(1): 87-96.

Cuong B.C. (2014). Picture fuzzy sets. Journal of Computer Science and Cybernetics. 30(4): 409- 420.

Cuong B.C. & Hai P.V. (2015). Some fuzzy logic operators for picture fuzzy sets. Seventh International Conference on Knowledge and Systems Engineering. pp. 132-137.

Dengfeng L. & Chuntian C. (2002). New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions. Pattern Recogn. Lett. 23: 221-225.

Dinh N.V. & Thao N.X. (2018). Some Measures of Picture Fuzzy Sets and Their Application in Multi-attribute Decision Making. I.J. Mathematical Sciences and Computing. 3: 23-41.

Dinh N.V., Thao N.X. & Chau N.M. (2015). On the picture fuzzy database: theories and application. Việt Nam National University of Agriculture, J. Sci. & Devel. 13(6): 1028-1035.

Dutta P. (2017). Medical Diagnosis via Distance Measures on Picture Fuzzy Sets. Amse Journals-Amseiieta publication. 54(2): 137-152.

Garg H. (2017). Some picture fuzzy aggregation operators and their applications to multicriteria decision-making. Arabian Journal for Science Engineering. 42(12): 5275-5290.

Hoa N.D., Son L.H. & Thong P.H. (2014). Weather nowcasting from satellite image sequences using the combination of picture fuzzy clustering and spatiotemporal regression. Proceeding of Conference of GISIDEAR. Đà Nẵng, Việt Nam.

Hoa N.D., Son L.H. & Thong P.H. (2016). Some improvements of fuzzy clustering algorithms using picture fuzzy sets and applications for geographic data clustering. VNU Journal of Science: Computer Science and Communication Engineering. 32(3): 32-38.

Hung W.L & Yang M.S. (2004). Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance. Pattern Recognition Letters. 25: 1603-1611.

Hung W.L. & Yang M.S. (2007). Similarity measures of intuitionistic fuzzy sets based on Lpmetric. International Journal of Approximate Reasoning. 46: 120-136.

Hwang C.M., Yang M.S. & Hung W.L. (2018). New similarity measures of intuitionistic fuzzy sets based on the Jaccard index with its application to clustering. International Journal of Intelligent Systems. 33(8): 1672-1688.

Joshi D. & Kumar S. (2018). An Approach to Multi-criteria Decision Making Problems Using Dice Similarity Measure for Picture Fuzzy Sets. Mathematics and Computing. Springer, Singapore.

Li D.F. & Cheng C. (2002). New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions. Pattern Recognition Letters. 23: 221-225.

Liu H.W. (2005). New similarity measures between intuitionistic fuzzy sets and between elements. Mathematical and Computer Modelling. 42: 61-70.

Mitchell H.B. (2003). On the Dengfeng-Chuntian similarity measure and its application to pattern recognition. Pattern Recognition Letters. 24: 3101-3104.

Peng X. & Dai J. (2017). Algorithm for picture fuzzy multiple attribute decision making based on new distance measure. International Journal for Uncertainty Quantification. 7: 177-187.

Phong P.H., Hieu D.T., Ngan R.T.H. & Them P.T. (2014). Some compositions of picture fuzzy relations, in: Proceedings of the 7th National Conference on Fundamental and Applied Information Technology Research, FAIR’7, Thai Nguyen. pp. 19-20.

Singh P. (2015). Correlation coefficients for picture fuzzy sets. Journal of Intelligent & Fuzzy Systems. 28(2): 591-604.

Son L.H. (2015). DPFCM: A novel distributed picture fuzzy clustering method on picture fuzzy sets. Expert System with Applications. 42(1): 51-66.

Son L.H. (2016). Generalized picture distance measure and applications to picture fuzzy clustering. Applied Soft Computing. 46: 248-295.

Son L.H. (2017). Measuring analogousness in picture fuzzy sets: from picture distance measures to picture association measures. Fuzzy Optimization and Decision Making. 16(3): 359-378.

Son L.H., Phong P.H. (2016). On the performance evaluation of intuitionistic vector similarity measures for medical diagnosis. Journal of Intelligent and Fuzzy Systems. 31: 1597-1608.

Son L.H., Viet P.V. & Hai P.V. (2017). Picture inference system: a new fuzzy inference system on picture fuzzy set. Applied Intelligence. 46(3): 652-669.

Szmidt E. (2014). Distances and Similarities in Intuitionistic Fuzzy Sets. Springer International Publishing Switzerland. p. 307.

Thao N.X. & Dinh N.V. (2015). Rough picture fuzzy set and picture fuzzy topologies. Journal of Computer Science and Cybernetics. 31(3): 245.

Thong P.H. & Son L.H. (2014). A new approach to multi-variable fuzzy forecasting using picture fuzzy clustering and picture fuzzy rule interpolation method. Knowledge and Systems Engineering, Springer, Cham. pp. 679-690.

Thong P.H. (2016). Picture fuzzy clustering: a new computational intelligence method. Soft Computing. 20(9): 3549-3562.

Thong P.H. (2016). Picture fuzzy clustering for complex data. Engineering Applications of Arti_cial Intelligence. 56: 121-130.

Thong P. H. (2017). Some novel hybrid forecast methods based on picture fuzzy clustering for weather now-casting from satellite image sequences. Applied Intelligence. 46(1):1-15.

Viet P.V., Chau H.T.M & Hai P.V. (2015). Some extensions of membership graphs for picture inference systems. Seventh International Conference on Knowledge and Systems Engineering, (KSE), IEEE. pp. 192-197.

Wei G.W. (2017). Some similarity measures for picture fuzzy sets and their applications to strategic decision making. Informatica. 28: 547-564.

Wei G.W. (2017). Some cosine similarity measures for picture fuzzy sets and their applications to strategic decision making. Informatica. 28(3): 547-564.

Wei G.W. (2018). Some similarity measures for picture fuzzy sets and their applications. Iranian Journal of Fuzzy Systems. 15(1): 77-89.

Wei G.W. (2018). The Generalized Dice Similarity Measures for Picture Fuzzy Sets and Their Applications. Informatica. 29(1): 107-124.

Xu Z.S., Chen J. & Wu J.J. (2008). Clustering algorithm for intuitionistic fuzzy sets. Information Sciences. 178: 3775-3790.

Xu Z.S. & Xia M.M. (2010). Some new similarity measures for intuitionistic fuzzy values and their application in group decision making. Journal of System Science and Engineering. 19: 430-452.

Ye J. (2011). Cosine similarity measures fot intuitionistic fuzzy sets and their applications. Mathematical and Computer Modelling. 53: 91-97.

Ye J. (2012). Multicriteria decision-making method using the Dice similarity measure based on the reduct intuitionistic fuzzy sets of interval-valued intuitionistic fuzzy sets. Applied Mathematical Modelling. 36: 4466-4472.

Ye J. (2016). Similarity measures of intuitionistic fuzzy sets based on cosine function for the decision making of mechanical design schemes. Journal of Intelligent and Fuzzy Systems. 30: 151-158.

Zadeh L.A. (1965). Fuzzy sets. Information and control. 8(3): 338-353.

Zeng S., Asharf S., Arif M. & Abdullah S. (2019). Application of Exponential Jensen Picture Fuzzy Divergence Measure in Multi-Criteria Group Decision Making. Mathematics. 7(2): 191-207.