

MỘT SỐ TÍNH CHẤT CỦA PHỦ SUY DẪN TỪ HỌ PHỦ TẬP THÔ

Nguyễn Đức Thuận^{a*}

^aKhoa Công nghệ Thông tin, Trường Đại học Nha Trang, Khánh Hòa, Việt Nam

Lịch sử bài báo

Nhận ngày 10 tháng 01 năm 2017 | Chính sửa ngày 30 tháng 04 năm 2017

Chấp nhận đăng ngày 07 tháng 07 năm 2017

Tóm tắt

Lý thuyết tập thô là một công cụ hiệu quả và cần thiết để xử lý tính mơ hồ và hạt trong hệ thống thông tin. Phủ dựa trên lý thuyết tập thô (phủ tập thô) được đề xuất như một sự tổng quát của lý thuyết tập thô cổ điển. Do phủ tập thô là tổng quát và phức tạp hơn, vì vậy cần thiết phát triển các cấu trúc mới, phức tạp nhằm phát hiện tính chất đặc trưng của nó. Trong bài báo này, chúng tôi nghiên cứu phủ suy dẫn từ họ phủ tập thô. Một số kết quả lý thuyết liên quan đến độ đo tính đặc trưng của tri thức được phát hiện.

Từ khóa: Đặc trưng tri thức; Họ phủ tập thô; Phủ suy dẫn; Tập thô.

1. ĐẶT VẤN ĐỀ

Trong nhiều ứng dụng thực tế, dữ liệu được tổ chức dưới dạng một phủ, thay cho các phân hoạch. Phủ được xây dựng trên tập thô xuất phát từ mối quan hệ dung sai thường được đề xuất, nghiên cứu nhằm xử lý các hệ thống thông tin thiếu dữ liệu. Với mục đích phát huy hiệu năng của phủ tập thô, nhiều tác giả đã nghiên cứu tính chất toán học, phát hiện các tính chất đặc trưng của phủ tập thô như: Rút gọn hệ thống thông tin và xây dựng hệ tiên đề cho phủ tập thô (William & Fei, 2002); Rút gọn tri thức và độ đo tri thức đặc trưng dựa vào phủ tập thô (Shi & Gong, 2008); Các phép xấp xỉ dựa vào phủ tập thô (Yao & Yao, 2012); Các độ đo cho phủ tập thô (Jianhua, Debiao, Huashi, & Haowei, 2014); và các toán tử láng giềng cho phủ tập thô (Lynn, Mauricio, Chriss, & Jonatan, 2016). Với mong muốn tìm được công cụ hiệu năng cao nhằm xử lý các hệ thống thông tin không đầy đủ, chúng tôi nhận thấy rằng phủ suy dẫn từ họ phủ tập thô được đề cập trong các bài báo của các tác giả Shi và Gong (2008); Shiping, Zhu, Quingxin và Fan (2012) có nhiều tiềm năng. Sự mở rộng các phủ suy dẫn này dựa trên các láng giềng các phần tử cũng cho

* Tác giả liên hệ: Email: thuan.inf@ntu.edu.vn

kết quả khả quan. Trong bài báo này chúng tôi khảo sát tính chất toán học của phủ suy dẫn, đề xuất một độ đo dùng để phân lớp dữ liệu trong các hệ thống thông tin quyết định không đầy đủ.

Cấu trúc của bài báo gồm bốn mục: Mục 1 đặt vấn đề; Mục 2 nêu một số khái niệm cơ sở; Mục 3 nêu một số kết quả đạt được. Cuối cùng là kết luận và hướng phát triển của bài báo.

2. MỘT SỐ KHÁI NIỆM CƠ SỞ

Định nghĩa 1: Phủ tập thô (William & Fei, 2002)

Cho U là tập vũ trụ, C là một họ các tập con khác rỗng của U . Nếu $\cup C = U$, C được gọi là một phủ của U . Cặp (U, C) gọi là một không gian xấp xỉ phủ.

Định nghĩa 2: Hệ thống láng giềng (Shi & Gong, 2008)

Cho C là 1 phủ của U , $x \in U$, hệ thống láng giềng của x , ký hiệu $C_{(C,x)}$,

$C_{(C,x)} = \{K \in C \mid x \in K\}$, viết gọn $C(x)$

Định nghĩa 3: Mô tả tối thiểu của x (William & Fei, 2002)

Cho (U, C) là một không gian xấp xỉ phủ, $x \in U$ họ tập hợp

$Md(x) = \{K \in C \mid x \in K \wedge (\forall S \in C \wedge x \in S \wedge S \subseteq K \Rightarrow K = S)\}$ được gọi là mô tả tối thiểu của x .

Định nghĩa 4: Láng giềng của x (Lynn và ctg., 2016).

Cho (U, C) là một không gian xấp xỉ phủ, $x \in U$, một láng giềng của x ký hiệu $N_C(x)$, xác định như sau: $N_C(x) = \cap \{K \in C \mid K \in Md(x)\}$

Mệnh đề 1: (Yao & Yao, 2012)

Cho (U, C) là một không gian xấp xỉ phủ, $x \in U$ thì $N_C(x) = \cap C_{(C,x)}$.

Định nghĩa 5: (Shi & Gong, 2008)

Cho $C = \{C_1, C_2, \dots, C_n\}$ là một phủ của U , $\forall x \in U$, $Cov(C) = \{\cap Md(x) \mid x \in U\}$ cũng là một phủ của U , chúng ta nói nó là một phủ suy dẫn của C .

Định nghĩa 6: (Shi & Gong, 2008)

Cho $\Delta = \{C_1, C_2, \dots, C_m\}$ là một họ phủ của U , $\forall x \in U$, đặt $\Delta_x = \cap\{\cap Md(x) \mid \cap Md(x) \in Cov(C_i), \forall i = \overline{1, m}\}$ thì $Cov(\Delta) = \{\Delta_x \mid x \in U\}$ cũng là một phủ của U , và chúng ta gọi là một phủ suy dẫn của Δ .

Nhận xét 1: Từ Định nghĩa 3 và Định nghĩa 6 ta có $\Delta_x = \bigcap_{c \in \Delta} N_c(x)$

Nếu Δ là một phân hoạch của U , thì $Cov(\Delta)$ cũng là một phân hoạch của U , khi đó Δ_x là một lớp tương đương chứa x .

Định nghĩa 7: Phủ mịn hơn (Jianhua và ctg., 2014)

Cho C_1, C_2 là 2 phủ của U . Nếu $\forall x \in U$, $C_1(x)$ và $C_2(x)$ thỏa:

$$(1) \quad \forall K_1 \in C_1(x), \exists K_2 \in C_2(x) : K_1 \subseteq K_2$$

$$(2) \quad \forall K_2 \in C_2(x), \exists K_1 \in C_1(x) : K_1 \subseteq K_2$$

Thì nói rằng C_1 mịn hơn C_2 , ký hiệu $C_1 \prec C_2$.

Định nghĩa 8: Xấp xỉ dưới (Shi & Gong, 2008)

Cho (U, C) là một không gian xấp xỉ phủ, $\forall X \subseteq U$, xấp xỉ dưới của X ứng với không gian xấp xỉ phủ (U, C) được xác định bởi (1).

$$\underline{C}(X) = \{x \in U \mid N_c(x) \subseteq X\} \tag{1}$$

Định nghĩa 9: (Shi & Gong, 2008)

Cho Δ là một họ phủ của U , P và Q là 2 phủ thuộc Δ . Miền dương của Q ứng với P được xác định $POS_p(Q) = \bigcup_{X \in Q} \underline{P}(X)$.

Định nghĩa 10: Mức đặc trưng (Shi & Gong, 2008)

Cho $S = (U, \Delta, F, D, G)$ là một hệ thống thông tin quyết định phủ, trong đó U là tập vũ trụ, Δ là một họ phủ của U , F là hàm thông tin, D thuộc tính quyết định, G là quan hệ tương đương phân hoạch theo thuộc tính quyết định D . Với mỗi $C \in \Delta$, mức đặc trưng của C đối với D được xác định như trong Công thức (2).

$$sig_C(D) = \frac{|POS_{Cov(\Delta)}(U/D)|}{|U|} - \frac{|POS_{Cov(\Delta-\{C\})}(U/D)|}{|U|} \quad (2)$$

Bổ đề 1: (Jianhua và ctg., 2014)

Cho C_1, C_2 là 2 phủ của U , nếu $C_1 \prec C_2$ thì $N_{C_1}(x) \subseteq N_{C_2}(x)$, $\forall x \in U_r$

Chứng minh: $\forall y \in N_{C_1}(x) \Rightarrow \forall K_1 \in C_1(x) : y \in K_1$, do $C_1 \prec C_2$ nên

$\forall K_2 \in C_2(x), \exists K_1 \in C_1(x) : K_1 \subseteq K_2$, do đó $y \in K_2, \forall K_2 \in C_2(x) \Rightarrow y \in \bigcap C_2(x)$

Nói khác hơn, $y \in N_{C_2}(x)$

Định nghĩa 11: Hệ thống thông tin quyết định không đầy đủ

Một hệ thống thông tin là một bộ bốn thành phần $S = \langle U, A, V, f \rangle$, trong đó: U là tập hữu hạn khác rỗng (tập vũ trụ); A là tập hữu hạn khác rỗng các thuộc tính; $V = \bigcup_{a \in A} V_a$, với $V_a = Dom(a) \neq \emptyset, |V_a| < \infty$; $f : U \times A \rightarrow V$, $f(x, a) = v \in V_a$; Nếu $\exists c \in A, x \in U$ mà $f(x, c)$ không xác định thì S gọi là hệ thống thông tin không đầy đủ hay hệ thống thông tin có thuộc tính thiếu dữ liệu; S được gọi là hệ thống thông tin quyết định nếu $A = C \cup D, C \cap D = \emptyset, C$ là tập thuộc tính điều kiện, D là tập thuộc tính quyết định. Hệ

thống thông tin quyết định ký hiệu thu gọn là $S = (U, C, D)$

3. MỘT SỐ KẾT QUẢ ĐẠT ĐƯỢC

Bổ đề 2: Cho $\Delta = \{C_1, C_2, \dots, C_m\}$ là một họ phủ của U , nếu $C_1 \prec C_2$ thì $\forall x \in U$:
 $(\Delta - \{C_1\})_x \supseteq (\Delta - \{C_2\})_x$

Chứng minh: Ký hiệu $\Delta'_x = \bigcap_{i=3}^m N_{C_i}(x)$. Ta có: $(\Delta \setminus \{C_1\})_x = N_{C_2}(x) \cap \Delta'_x$;
 $(\Delta \setminus \{C_2\})_x = N_{C_1}(x) \cap \Delta'_x$. Từ Bổ đề 1, nếu $C_1 \prec C_2$ thì $N_{C_1}(x) \subseteq N_{C_2}(x)$ nên
 $(\Delta - \{C_1\})_x \supseteq (\Delta - \{C_2\})_x$

Mệnh đề 1: Cho $\Delta = \{C_1, C_2, \dots, C_m\}$ là một họ phủ của U , nếu $C_1 \prec C_2$ thì:

$$(1) \text{Cov}(\Delta - \{C_2\}) \succ \text{Cov}(\Delta - \{C_1\});$$

$$(2) \text{POS}_{\text{Cov}(\Delta - \{C_2\})}(U/D) \supseteq \text{POS}_{\text{Cov}(\Delta - \{C_1\})}(U/D).$$

Chứng minh:

(1) Là kết quả trực tiếp từ Bổ đề 2, Định nghĩa 6 và Định nghĩa 7;

(2) $\forall x \in \text{POS}_{\text{Cov}(\Delta - \{C_1\})}(U/D) \Rightarrow x \in \bigcup_{X \in U/D} \text{Cov}(\Delta - \{C_1\})(X)$, mà

$$(\Delta - \{C_1\})_x \subseteq X \Rightarrow (\Delta - \{C_2\})_x \subseteq X \text{ do } (\Delta - \{C_2\})_x \subseteq (\Delta - \{C_1\})_x \text{ (bổ đề 2)}$$

Vì vậy, $x \in \text{POS}_{\text{Cov}(\Delta - \{C_2\})}(U/D)$. Nói khác hơn,

$$\text{POS}_{\text{Cov}(\Delta - \{C_2\})}(U/D) \supseteq \text{POS}_{\text{Cov}(\Delta - \{C_1\})}(U/D).$$

Mệnh đề 2:

Cho $S = (U, \Delta, F, D, G)$ là một hệ thống thông tin quyết định phủ. Nếu $C_1 \prec C_2$ thì

$$\text{sig}_{C_1}(D) \geq \text{sig}_{C_2}(D)$$

Chứng minh: là kết quả trực tiếp từ Định nghĩa 10, kết quả (b) của Mệnh đề 1.

Nhận xét 2: Bài toán phân lớp, xây dựng cây quyết định, rút trích luật quyết định là bài toán quan trọng trong lĩnh vực khai phá dữ liệu. Đối với các hệ thống thông tin quyết định không đầy đủ cũng có nhiều tác giả quan tâm và đưa ra các giải pháp tiếp cận bài toán khác nhau (Marzena, 1998; Li, Li, & Wu, 2009). Các giải pháp này thường điểm mấu chốt là tìm các thuộc tính phân chia các đối tượng vào các lớp xấp xỉ các lớp quyết định nhanh nhất. Tư tưởng chung trong việc lựa chọn thuộc tính phân chia là sử dụng thuộc tính xấp xỉ phân lớp tốt nhất tập thuộc tính quyết định trên tập đối tượng cần phân chia. Để làm điều này, trong hệ thống thông tin đầy đủ một số tác giả đã đề xuất các độ đo để xác định thuộc tính phân chia (Nguyễn, 2010). Trong hệ thống thông tin không đầy đủ, từ tính chất ở Mệnh đề 2 cho thấy thuộc tính phân lớp mịn hơn thì mức đặc trưng cao hơn, có nghĩa là sự phân lớp của thuộc tính này xấp xỉ phân lớp của tập thuộc tính quyết định là tốt nhất, vì vậy có thể sử dụng làm độ đo để xác định thuộc tính phân chia các đối tượng. Để minh họa độ đo mức đặc trưng trên một hệ thống thông tin không đầy đủ, chúng ta xem xét ví dụ như sau:

Ví dụ: Xét hệ thống thông tin quyết định không đầy đủ $S=(U, AT, D)$ như trong Bảng 1. Trong đó, $U = \{u_1, u_2, u_3, u_4, u_5, u_6\}$ là tập vũ trụ; $AT=\{Price, Mileage, Size, Max_Speed\}$ là tập thuộc tính điều kiện; D thuộc tính quyết định.

Bảng 1. Ví dụ về hệ thống thông tin quyết định không đầy đủ $S = (U, AT, D)$

U	Price (p)	Mileage (M)	Size (S)	Max_speed (A)	D
u1	High	High	Full	Low	Good
u2	Low	*	Full	Low	Good
u3	*	*	Compact	High	Poor
u4	High	*	Full	High	Good
u5	*	*	Full	High	Excel
u6	Low	High	Full	*	Good

Xét quan hệ dung sai trên mỗi thuộc tính như sau:

$$\forall A \in AT, SIM(A) = \{(x,y) \in U \times U : A(x)=A(y) \text{ or } A(x)=* \text{ or } A(y)=*\}$$

Ta có các phủ của U :

$$C_1 = C_P = \{\{u_1, u_3, u_4, u_5\}, \{u_2, u_3, u_5, u_6\}\}$$

$$C_2 = C_M = \{\{u_1, u_2, u_3, u_4, u_5, u_6\}\}$$

$$C_3 = C_S = \{\{u_1, u_2, u_4, u_5, u_6\}, \{u_3\}\}$$

$$C_4 = C_A = \{\{u_1, u_2, u_6\}, \{u_3, u_4, u_5, u_6\}\}$$

$$\Delta = \{C_1, C_2, C_3, C_4\}; U/D = \{\{u_1, u_2, u_4, u_6\}, \{u_3\}, \{u_5\}\}$$

$$\Delta_{u_1} = \{u_1\}; \Delta_{u_2} = \{u_2, u_6\}; \Delta_{u_3} = \{u_3\}; \Delta_{u_4} = \{u_4, u_5\}; \Delta_{u_5} = \{u_5\}; \Delta_{u_6} = \{u_6\}$$

$$(\Delta - \{C_1\}):$$

$$(\Delta - \{C_1\})_{u_1} = \{u_1, u_2, u_6\}; (\Delta - \{C_1\})_{u_2} = \{u_1, u_2, u_6\}; (\Delta - \{C_1\})_{u_3} = \{u_3\};$$

$$(\Delta - \{C_1\})_{u_4} = \{u_4, u_5, u_6\}; (\Delta - \{C_1\})_{u_5} = \{u_4, u_5, u_6\}; (\Delta - \{C_1\})_{u_6} = \{u_6\};$$

$$(\Delta - \{C_2\}):$$

$$(\Delta - \{C_2\})_{u_1} = \{u_1\}; (\Delta - \{C_2\})_{u_2} = \{u_2, u_6\}; (\Delta - \{C_2\})_{u_3} = \{u_3\};$$

$$(\Delta - \{C_2\})_{u_4} = \{u_4, u_5\}; (\Delta - \{C_2\})_{u_5} = \{u_5\}; (\Delta - \{C_2\})_{u_6} = \{u_6\};$$

$$(\Delta - \{C_3\}):$$

$$(\Delta - \{C_3\})_{u_1} = \{u_1\}; (\Delta - \{C_3\})_{u_2} = \{u_2, u_6\}; (\Delta - \{C_3\})_{u_3} = \{u_3, u_5\};$$

$$(\Delta - \{C_3\})_{u_4} = \{u_3, u_4, u_5\}; (\Delta - \{C_3\})_{u_5} = \{u_3, u_5\}; (\Delta - \{C_3\})_{u_6} = \{u_6\};$$

$$(\Delta - \{C_4\}):$$

$$(\Delta - \{C_4\})_{u_1} = \{u_1, u_4, u_5\}; (\Delta - \{C_4\})_{u_2} = \{u_2, u_5, u_6\}; (\Delta - \{C_4\})_{u_3} = \{u_3\};$$

$$(\Delta - \{C_4\})_{u_4} = \{u_1, u_4, u_5\}; (\Delta - \{C_4\})_{u_5} = \{u_5\}; (\Delta - \{C_4\})_{u_6} = \{u_2, u_5, u_6\};$$

$$POS_{Cov(\Delta - \{C_1\})}(U/D) = \{u_1, u_2, u_3, u_6\}; POS_{Cov(\Delta - \{C_2\})}(U/D) = \{u_1, u_2, u_3, u_5, u_6\};$$

$$POS_{Cov(\Delta - \{C_3\})}(U/D) = \{u_1, u_2, u_6\}; POS_{Cov(\Delta - \{C_4\})}(U/D) = \{u_3, u_5\};$$

$$POS_{Cov(\Delta)}(U/D) = \{u_1, u_2, u_3, u_5, u_6\};$$

$$sig_{C_1}(D) = \frac{1}{6}; sig_{C_2}(D) = 0; sig_{C_3}(D) = \frac{1}{3}; sig_{C_4}(D) = \frac{1}{2}$$

Chọn thuộc tính phân chia tập đối tượng là Max_Speed (A), U được phân thành hai lớp:

$$\text{Max_Speed (A)= Low} \equiv \{u_1, u_2, u_6\} \subset D_{Good} = \{u_1, u_2, u_4, u_6\}$$

$$\text{Max_Speed (A)= High} \equiv \{u_3, u_4, u_5, u_6\}$$

Xét hệ thống thông tin quyết định $S' = \{U_1, AT_1, D\}$, với $U_1 = \{u_3, u_4, u_5, u_6\}$, $AT_1 = \{\text{Price (P), Mileage (M), Size (S)}\}$

Ta có các phủ của U_1 :

$$C'_1 = C'_P = \{\{u_3, u_4, u_5\}, \{u_3, u_5, u_6\}\}; C'_2 = C'_M = \{\{u_3, u_4, u_5, u_6\}\}$$

$$C'_3 = C'_S = \{\{u_3\}, \{u_4, u_5, u_6\}\}; \Delta' = \{C'_1, C'_2, C'_3\}; U_1/D = \{\{u_3\}, \{u_4, u_6\}, \{u_5\}\}$$

$$\Delta'_{u_3} = \{u_3\}; \Delta'_{u_4} = \{u_4, u_5\}; \Delta'_{u_5} = \{u_5\}; \Delta'_{u_6} = \{u_5, u_6\}$$

$$(\Delta' - \{C'_1\})_{u_3} = \{u_3\}; (\Delta' - \{C'_1\})_{u_4} = \{u_4, u_5, u_6\}; (\Delta' - \{C'_1\})_{u_5} = \{u_4, u_5, u_6\};$$

$$(\Delta' - \{C'_1\})_{u_6} = \{u_4, u_5, u_6\};$$

$$(\Delta' - \{C'_2\})_{u_3} = \{u_3\}; (\Delta' - \{C'_2\})_{u_4} = \{u_4, u_5\}; (\Delta' - \{C'_2\})_{u_5} = \{u_5\};$$

$$(\Delta' - \{C'_2\})_{u_6} = \{u_5, u_6\};$$

$$(\Delta' - \{C_3'\})_{u_3} = \{u_3, u_5\}; \quad (\Delta' - \{C_3'\})_{u_4} = \{u_3, u_4, u_5\}; \quad (\Delta' - \{C_3'\})_{u_5} = \{u_3, u_5\};$$

$$(\Delta' - \{C_3'\})_{u_6} = \{u_3, u_5, u_6\};$$

$$POS_{Cov(\Delta' - \{C_1'\})}(U_1 / D) = \{u_3\}; \quad POS_{Cov(\Delta' - \{C_2'\})}(U_1 / D) = \{u_3, u_5\}$$

$$POS_{Cov(\Delta' - \{C_3'\})}(U_1 / D) = \emptyset \Rightarrow sig_{C_1'}(D) = \frac{1}{4}; sig_{C_2'}(D) = 0; sig_{C_3'}(D) = \frac{1}{2}$$

Chọn thuộc tính phân chia tập đối tượng là Size (S), U_1 được phân thành 2 lớp:

$$\text{Size (S)= Compact} \equiv \{u_3\} \equiv D_{Poor} = \{u_3\}; \quad \text{Size (S)= Full} \equiv \{u_4, u_5, u_6\}$$

Xét hệ thống thông tin quyết định $S' = \{U_2, AT_2, D\}$, với $U_2 = \{u_4, u_5, u_6\}$, $AT_2 = \{Price (P), Mileage (M)\}$. Ta có các phủ của U_2 :

$$C_1'' = C_P'' = \{\{u_4, u_5\}, \{u_5, u_6\}\}; \quad C_2'' = C_M'' = \{\{u_4, u_5, u_6\}\}$$

$$U_2 / D = \{\{u_4, u_6\}, \{u_5\}\}$$

$$\Delta'' = \{C_1'', C_2''\}; \quad \Delta''_{u_4} = \{u_4, u_5\}; \quad \Delta''_{u_5} = \{u_5\}; \quad \Delta''_{u_6} = \{u_5, u_6\};$$

$$(\Delta'' - \{C_1''\})_{u_4} = (\Delta'' - \{C_1''\})_{u_5} = (\Delta'' - \{C_1''\})_{u_6} = \{u_4, u_5, u_6\}$$

$$(\Delta'' - \{C_2''\})_{u_4} = \{u_4, u_5\}; \quad (\Delta'' - \{C_2''\})_{u_5} = \{u_5\}; \quad (\Delta'' - \{C_2''\})_{u_6} = \{u_5, u_6\}$$

$$POS_{Cov(\Delta'' - \{C_1''\})}(U_2 / D) = \emptyset; \quad POS_{Cov(\Delta'' - \{C_2''\})}(U_2 / D) = \{u_5\} = POS_{Cov(\Delta'')}(U_2 / D)$$

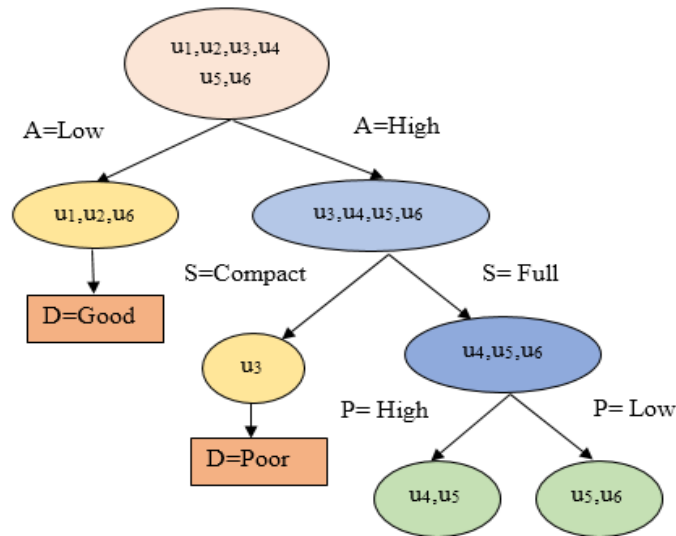
$$sig_{C_1''}(D) = \frac{1}{3}; \quad sig_{C_2''}(D) = 0 \Rightarrow \text{Chọn thuộc tính Price (P) để phân lớp, } U_2 \text{ phân}$$

thành Price=High $\equiv \{u_4, u_5\}$; Price=Low $\equiv \{u_5, u_6\}$; và không còn phân chia được nữa.

Như vậy ta xây dựng được cây quyết định biểu diễn bởi Hình 1, suy dẫn được 2 luật quyết định tương ứng:

$$(1) \text{Max_Speed} = \text{Low} \rightarrow D = \text{Good}$$

$$(2) Max_Speed = High \wedge Size = Compact \rightarrow D = Poor$$



Hình 1. Cây quyết định

4. KẾT LUẬN

Trong bài báo này, chúng tôi phát hiện được một số tính chất toán học về phủ suy dẫn từ họ phủ tập thô liên quan đến mức đặc trưng của phủ đối với họ phủ, và mối quan hệ giữa láng giềng của một đối tượng x ứng với 2 phủ có mối quan hệ mịn hơn. Những kết quả này có thể dùng để phát triển thành độ đo hỗ trợ xây dựng cây quyết định hay rút trích đặc trưng. Trong thời gian đến, chúng tôi sẽ thử nghiệm trên các bộ dữ liệu UCI để đánh giá cụ thể hơn.

TÀI LIỆU THAM KHẢO

- Jianhua, D., Debiao, H., Huashi, S., & Haowei, T. (2014). Uncertainty measurement for covering rough sets. *International Journal of Uncertainty, Fuzziness and Knowledge-Base Systems*, 22(2), 217-233.
- Li, P., Li, F., & Wu, Q. (2009). *Uncertainty analysis to the decision rule of incomplete information system*. Paper presented at The Information Technology and Applications Conference, USA.
- Lynn, D., Mauricio, R., Chriss, C., & Jonatan, G. (2016). Neighborhood operator for covering-based rough sets. *International Sciences*, 336, 21-24.
- Marzena, K. (1998). Rough set approach to incomplete information system. *Information Sciences*, 112, 39-49.

- Mauricio, R., Chris, C., & Jonatan, G. (2013). *Characterization of neighborhood operators for covering based rough sets using duality and adroitness*. Paper presented at The Eureka-2013, Fourth International Workshop Conference, USA.
- Nguyễn, T. T. (2010). Về một metric trên họ các phân hoạch của một tập hữu hạn. *Tạp chí Tin học và Điều khiển học*, 20(1), 75-87.
- Shi, Z., & Gong, Z. (2008). Knowledge reduction and knowledge significance measure based on covering rough sets. *International Journal of Pure and Applied Mathematics*, 48(1), 1-9.
- Shiping, W., Zhu, W., Qingxin, Z., & Fan, M. (2012). Covering base. *Journal of Information and Computer Sciences*, 9(5), 1343-1355.
- William, Z., & Fei, Y. W. (2002). Reduction and axiomatization of covering generalized rough sets. *Information Sciences*, 152, 217-230.
- Yao, Y. Y., & Yao, B. (2012). Covering based rough sets approximation. *Information Sciences*, 200, 91-107.

SOME PROPERTIES OF INDUCED COVER FROM FAMILY OF COVERING ROUGH SETS

Nguyễn Đức Thuận^{a*}

^a*The Faculty of Information Technology, Nhatrang University, Khanhhoa, Vietnam*

^{*}*Corresponding author: Email: thuan.inf@ntu.edu.vn*

Article history

Received: January 10th, 2017 | Received in revised form: April 30th, 2017

Accepted: July 07th, 2017

Abstract

Rough set theory is an efficient and essential tool for dealing with vagueness and granularity in information systems. Covering based rough set theory is proposed as a significant generalization of classical rough sets. It is more general and complex than classical rough set theory, hence there is much need to develop sophisticated structures to characterize covering-based rough sets. In this paper, we studied induced cover from family of covering rough set. Some theoretical results related to the measuring significance of knowledge were obtained and investigated.

Keywords: Family of covering rough set; Induced cover; Knowledge significance; Rough set.
