

MỘT CẢI TIẾN TRONG ĐÁNH GIÁ ĐỘ TƯƠNG TỰ NGỮ NGHĨA GIỮA HAI KHÁI NIỆM TRONG KỸ NGHỆ ONTOLOGY

Hoàng Hữu Hạnh^{*}, Nguyễn Văn Trung⁺
Học Viện Công nghệ Bưu chính Viễn thông
⁺ Trường Đại học Khoa học Huế

Tóm tắt: Đánh giá độ tương tự giữa hai khái niệm trong ontology đóng vai trò quan trọng trong các bài toán thuộc lĩnh vực kỹ nghệ ontology bao gồm: đối sánh, so khớp, tích hợp ontology. Bài báo này nêu một phương pháp mới để đánh giá độ tương tự ngữ nghĩa giữa hai khái niệm trong ontology. Bài báo cũng trình bày phương pháp quy hoạch động để tính độ tương tự ngữ nghĩa giữa hai biểu thức khái niệm bất kỳ dựa vào ontology cho trước.

Từ khóa: Độ tương tự ngữ nghĩa, ngữ nghĩa, ontology, Web ngữ nghĩa

I. GIỚI THIỆU

Web ngữ nghĩa là một lĩnh vực nghiên cứu đang phát triển nhanh trong phát triển của trí tuệ nhân tạo và các hệ thống tri thức; và nhận được sự quan tâm của cộng đồng nghiên cứu trong thập niên vừa qua. Công nghệ Web ngữ nghĩa đang được áp dụng vào nhiều lĩnh vực khác nhau trong thực tế như tin-sinh học, tin học trong y tế, quản trị tri thức, công nghệ phần mềm, xử lý ngôn ngữ tự nhiên, v.v... Thành phần quan trọng trong cấu trúc của Web ngữ nghĩa và các ứng dụng của nó là ontology.

Trong một ontology người ta định nghĩa các thực thể -- bao gồm khái niệm, thuộc tính, cá thể -- và mối quan hệ giữa các thực thể này theo ngữ nghĩa được quy định tường minh bởi một ngôn ngữ logic xác định. Ngôn ngữ ontology được sử dụng phổ biến nhất hiện nay là OWL 2², được chuẩn hoá bởi tổ chức W3C vào năm 2012. Mối quan hệ thường được xét đến nhiều nhất giữa các thực thể trong một ontology chính là mối quan hệ phân cấp, theo đó, các thực thể trong ontology được xếp trên các cây phân cấp: cây phân cấp khái niệm và cây phân cấp thuộc tính.

Một trong những khâu quan trọng trong các bài toán thuộc lĩnh vực kỹ nghệ ontology như đối sánh ontology, so khớp ontology, tích hợp ontology chính là đánh giá mức độ tương tự giữa hai thực thể của một hoặc nhiều ontology.

Trên thực tế, do tính tương tự của khái niệm hay thuộc tính khi xét trên cây phân cấp của chúng nên các kỹ thuật đánh giá độ tương tự giữa hai khái niệm cũng có thể được áp dụng cho hai thuộc tính. Theo T. Slimani [2], các kỹ thuật đánh giá độ tương tự giữa hai khái niệm trong ontology được phân thành các loại sau đây:

- (1) Đánh giá dựa vào cấu trúc: còn được gọi là đánh giá dựa vào việc đếm số cạnh nối giữa hai khái niệm trên cây phân cấp khái niệm.
- (2) Đánh giá dựa vào nội dung thông tin: đánh giá độ tương tự giữa hai khái niệm dựa vào tần suất xuất hiện của các từ khoá tương ứng với hai khái niệm trong một tập hợp tài liệu cho trước.
- (3) Đánh giá dựa vào đặc trưng của hai khái niệm đang xét. Đặc trưng của khái niệm có thể là tập hợp các từ đồng nghĩa của nó trong từ điển WordNet hoặc tập hợp mối quan hệ của khái niệm trên cấu trúc phân cấp.

Ngoài ba nhóm phương pháp trên, trong một số trường hợp, người ta cũng sử dụng kết hợp phương pháp dựa vào cấu trúc với phương pháp đánh giá dựa vào nội dung thông tin hoặc đánh giá dựa vào đặc trưng của khái niệm.

Nhóm phương pháp đánh giá dựa vào cấu trúc được nhận xét là đơn giản do chỉ phụ thuộc cấu trúc phân cấp của ontology chứ không cần tham khảo đến nguồn dữ liệu ngoài như hai nhóm phương pháp còn lại. Một trong những phương pháp đánh giá độ tương tự giữa hai khái niệm thuộc nhóm này được ra đời sớm nhất là phương pháp của Wu và Palmer [4]. Mặc dù được ra đời sớm và được nhận xét là đơn giản nhưng phương pháp đánh giá của Wu và Palmer vẫn còn được sử dụng trong những năm gần đây, trong nhiều công trình thuộc các lĩnh vực Web ngữ nghĩa hay xử lý ngôn ngữ tự nhiên [5].

Bài báo này sẽ phân tích và đề xuất một cải tiến cho phương pháp Wu và Palmer để đánh giá độ tương tự giữa hai khái niệm. Các phần tiếp theo của bài báo được trình bày như sau: Phần 2 trình bày các khái niệm cơ bản để tính độ tương tự ngữ nghĩa giữa hai khái niệm trên cây phân cấp của ontology. Phần 3 – là đóng góp chính của bài báo – nêu các điểm còn tồn tại và đề xuất cải tiến cho phương pháp đánh giá của Wu-Palmer. Trong phần này, bài báo cũng trình bày phương pháp quy hoạch động để đánh giá độ

Tác giả liên lạc: Hoàng Hữu Hạnh,
Email: hoanghuuhanh@ptit.edu.vn
Đến tòa soạn: 11 /2019, chỉnh sửa: 12 /2019, chấp nhận đăng: 12/2019.

tương tự ngữ nghĩa giữa hai biểu thức khái niệm. Phần 4 của bài báo nêu kết luận và hướng mở rộng của bài báo..

II. PHƯƠNG PHÁP WU-PALMER ĐÁNH GIÁ ĐỘ TƯƠNG TỰ GIỮA HAI KHÁI NIỆM TRONG ONTOLOGY

Gọi \mathcal{O} là một ontology. $C_{\mathcal{O}}$ là tập các khái niệm có tên trong \mathcal{O} . Với hai khái niệm $c_1, c_2 \in C_{\mathcal{O}}$, chúng ta ký hiệu:

- $DP_{\mathcal{O}}(c_1)$ là tập các khái niệm cha trực tiếp của c_1 trên \mathcal{O} .
- $DC_{\mathcal{O}}(c_1)$ là tập các khái niệm con trực tiếp của c_1 trên \mathcal{O} .
- $LCP_{\mathcal{O}}(c_1, c_2)$ là tập các khái niệm cha chung nhỏ nhất của c_1 và c_2 trên \mathcal{O} .

Lưu ý rằng, chúng ta luôn ngầm định, khái niệm đỉnh \top và khái niệm đáy \perp là hai khái niệm có tên trong \mathcal{O} , tức là: $\top \in C_{\mathcal{O}}$ và $\perp \in C_{\mathcal{O}}$.

Chúng ta định nghĩa số cạnh nối giữa hai khái niệm trên cây phân cấp khái niệm của ontology \mathcal{O} như sau:

Định nghĩa 1. (Số cạnh nối giữa hai khái niệm trên cây phân cấp của ontology) *Gọi \mathcal{O} là một ontology. $C_{\mathcal{O}}$ là tập các khái niệm có tên trong \mathcal{O} . Ta nói $k \in \mathbb{N}$ là số cạnh nối giữa hai khái niệm $c_1, c_2 \in C_{\mathcal{O}}$, ký hiệu là $\mu_{\mathcal{O}}(c_1, c_2) = k$ nếu tồn tại d_1, d_2, \dots, d_k là dãy ngắn nhất các khái niệm có tên trong \mathcal{O} sao cho:*

$$\begin{cases} c_1 \equiv d_1 \\ c_2 \equiv d_2 \\ d_i \in DC_{\mathcal{O}}(d_{i+1}) \forall i = 1, 2, \dots, k - 1 \end{cases}$$

Ta quy ước:

- $\mu_{\mathcal{O}}(c, c) = 0$ với $\forall c \in C_{\mathcal{O}}$;
- $\mu_{\mathcal{O}}(c_1, c_2) = +\infty$ nếu $c_1 \equiv \perp$ hoặc c_1 không phải là khái niệm con của c_2 .
- Trong trường hợp ontology \mathcal{O} đã được xác định rõ, chúng ta có thể bỏ qua chỉ số \mathcal{O} để viết μ thay vì $\mu_{\mathcal{O}}$.

Phương pháp đánh giá độ tương tự ngữ nghĩa giữa hai khái niệm của Wu và Palmer được dựa trên số cạnh nối giữa hai khái niệm này đến khái niệm cha chung nhỏ nhất của chúng trên cây phân cấp khái niệm.

Định nghĩa 2. *Gọi \mathcal{O} là một ontology. $C_{\mathcal{O}}$ là tập các khái niệm có tên trong \mathcal{O} . Độ tương tự ngữ nghĩa giữa hai khái niệm $c_1, c_2 \in C_{\mathcal{O}}$, ký hiệu là $\sigma_{\mathcal{O}}(c_1, c_2)$ được xác định như sau:*

$$\sigma_{\mathcal{O}}(c_1, c_2) = \min \left\{ \frac{2 \cdot \mu_{\mathcal{O}}(c_0, \top)}{\mu_{\mathcal{O}}(c_1, c_0) + \mu_{\mathcal{O}}(c_2, c_0) + 2 \cdot \mu_{\mathcal{O}}(c_0, \top)} \mid c_0 \in LCP_{\mathcal{O}}(c_1, c_2) \right\}$$

Chúng ta có một số nhận xét rút ra trực tiếp từ Định nghĩa 1 và Định nghĩa 2 như mệnh đề dưới đây:

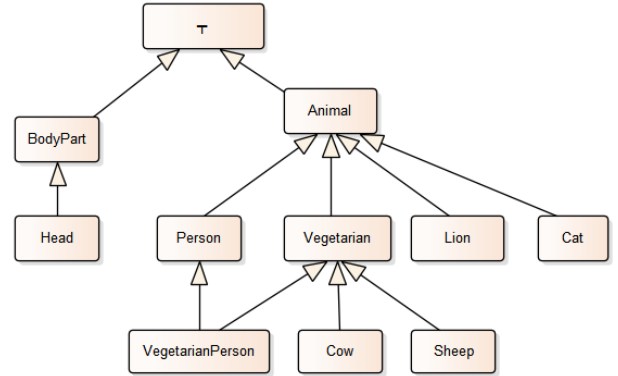
Mệnh đề 1. *Gọi \mathcal{O} là một ontology. $C_{\mathcal{O}}$ là tập các khái*

niệm có tên trong \mathcal{O} . Với mọi khái niệm $c_1, c_2 \in C_{\mathcal{O}}$, ta có:

- $\sigma(c_1, c_2) = \sigma(c_2, c_1)$
- $0 \leq \sigma(c_1, c_2) \leq 1$
- $\sigma(c_1, c_2) = 1$ khi $c_1 \equiv c_2$
- $\sigma(c_1, c_2) = 0$ khi $LCP_{\mathcal{O}}(c_1, c_2) = \{\top\}$ hoặc $c_1 \equiv \perp$ hoặc $c_2 \equiv \perp$

Dưới đây là ví dụ cho thấy các đặc điểm của phương pháp tính độ tương tự ngữ nghĩa giữa hai khái niệm trong ontology.

Ví dụ 1. Cho ontology \mathcal{O} với các khái niệm lập thành cây phân cấp như hình dưới đây:



Hình 1. Cây phân cấp khái niệm của ontology động vật

Độ tương tự giữa hai khái niệm *Lion* và *Cat* trong ontology \mathcal{O} được xác định lần lượt theo từng bước như sau:

- $LCP_{\mathcal{O}}(Lion, Cat) = \{Animal\}$
- $\mu(Animal, \top) = 1$
- $\mu(Lion, Animal) = 1$
- $\mu(Cat, Animal) = 1$

$$\sigma(Lion, Cat) = \frac{2 \cdot \mu(Animal, \top)}{\mu(Lion, Animal) + \mu(Cat, Animal) + 2 \cdot \mu(Animal, \top)} = \frac{2 \cdot 1}{1 + 1 + 2 \cdot 1} = 0.5$$

Tương tự như vậy, độ tương tự ngữ nghĩa giữa hai khái niệm *Sheep* và *Cat* – cũng có khái niệm cha chung nhỏ nhất là *Animal* được tính như sau:

$$\sigma(Sheep, Cat) = \frac{2 \cdot \mu(Animal, \top)}{\mu(Sheep, Animal) + \mu(Cat, Animal) + 2 \cdot \mu(Animal, \top)} = \frac{2 \cdot 1}{2 + 1 + 2 \cdot 1} \approx 0.3$$

3

Ở Ví dụ 1, chúng ta có thể thấy rằng, $\sigma(Lion, Cat) > \sigma(Sheep, Cat)$. Điều phản ánh đúng hình ảnh trực quan của cây phân cấp khái niệm ở Hình 1: Hai

khái niệm càng cách xa khái niệm cha chung nhỏ nhất thì có độ giống nhau càng thấp. Đây cũng chính là ưu điểm của phương pháp đánh giá Wu-Palmer.

Tuy vậy, phương pháp đánh giá Wu-Palmer có một nhược điểm: không thể đánh giá được mức độ giống nhau của các cặp khái niệm khi chúng có khái niệm cha chung nhỏ nhất là T. Điều này được chỉ ra ở Ví dụ 2 dưới đây:

Ví dụ 2. Xét ontology ở Ví dụ 1. Ta có:

$$LCP_O(BodyPart, Cat) =$$

$$LCP_O(BodyPart, Animal) = \{T\}$$

$$\sigma(BodyPart, Cat) =$$

$$\frac{2 \cdot \mu(T, T)}{\mu(BodyPart, T) + \mu(Cat, T) + 2 \cdot \mu(T, T)} = 0$$

$$\sigma(BodyPart, Animal) =$$

$$\frac{2 \cdot \mu(T, T)}{\mu(BodyPart, T) + \mu(Animal, T) + 2 \cdot \mu(T, T)} = 0$$

Như vậy, theo cách tính của Wu và Palmer, cả hai cặp khái niệm $(BodyPart, Cat)$ và $(BodyPart, Animal)$ đều có độ tương tự như nhau, và bằng 0.

Chúng ta cần phân biệt mức độ tương tự nhau giữa hai cặp khái niệm $(BodyPart, Cat)$ và $(BodyPart, Animal)$ thay vì đánh giá chúng bằng nhau (và đều bằng 0) như cách đánh giá của Wu và Palmer. Điều này thật sự cần thiết trong các bài toán lựa chọn khái niệm thuộc về hai nhánh gốc của cây phân cấp khái niệm, chẳng hạn như bài toán xử lý xung đột mức khái niệm trong quá trình tích hợp ontology theo cách tiếp cận của lý thuyết đồng thuận [3]. Phần tiếp theo của bài báo sẽ nêu những điều chỉnh để khắc phục nhược điểm này của phương pháp Wu-Palmer.

III. CÁC ĐỀ XUẤT ĐỂ KHẮC PHỤC HẠN CHẾ CỦA PHƯƠNG PHÁP WU-PALMER

Phần này của bài báo trình bày cải tiến phương pháp Wu-Palmer để tính độ tương tự ngữ nghĩa giữa hai khái niệm trong ontology theo các tiêu chí:

- Vẫn đảm bảo tính chất của hàm tính độ tương tự như đã chỉ ra ở Mệnh đề 1.
- Giữ được đặc tính trực quan của cây phân cấp khái niệm như phương pháp tính độ tương tự ngữ nghĩa của Wu-Palmer: Hai khái niệm càng cách xa khái niệm cha chung nhỏ nhất thì có độ giống nhau càng thấp.
- Phân biệt được mức độ giống nhau giữa các cặp khái niệm có khái niệm cha chung nhỏ nhất là *khái niệm đỉnh* T.

Trên thực tế, chúng ta chỉ cần điều chỉnh cách tính “số cạnh nối giữa hai khái niệm trên cây phân cấp của ontology” ở Định nghĩa 1 như sau. “Số cạnh” này được gọi là số cạnh μ' trong bài báo này.

Định nghĩa 3. (Số cạnh nối μ' giữa hai khái niệm trên cây phân cấp của ontology) Gọi O là một ontology. C_O là tập các khái niệm có tên trong O . Số cạnh nối trực tiếp μ' giữa hai khái niệm $c_1, c_2 \in C_O$, ký hiệu là $\mu'_O(c_1, c_2)$ được xác định như sau:

$$\mu'_O(c_1, c_2) = \begin{cases} \mu_O(c_1, c_2) & \text{với } c_2 \neq T \\ \mu_O(c_1, c_2) + 1 & \text{với } c_2 \equiv T \end{cases}$$

Trong trường hợp ontology O đã được xác định rõ, chúng ta có thể bỏ qua chỉ số O để viết μ' thay vì μ'_O .

Bằng cách sử dụng μ' thay cho μ trong công thức ở Định nghĩa 2, chúng ta sẽ có công thức mới để đánh giá độ tương tự ngữ nghĩa giữa hai khái niệm như sau:

Định nghĩa 4. Gọi O là một ontology. C_O là tập các khái niệm có tên trong O . Độ tương tự ngữ nghĩa σ' giữa hai khái niệm $c_1, c_2 \in C_O$, ký hiệu là $\sigma'_O(c_1, c_2)$ được xác định như sau:

$$\sigma'_O(c_1, c_2) = \min \left\{ \frac{2 \cdot \mu'_O(c_0, T)}{\mu'_O(c_1, c_0) + \mu'_O(c_2, c_0) + 2 \cdot \mu'_O(c_0, T)} \mid c_0 \in LCP_O(c_1, c_2) \right\}$$

Có thể thấy rằng, σ'_O cũng có các tính chất tương tự như σ_O được nêu ở Mệnh đề 1. Tức là:

- $\sigma'(c_1, c_2) = \sigma'(c_2, c_1)$
- $0 \leq \sigma'(c_1, c_2) \leq 1$
- $\sigma'(c_1, c_2) = 1$ khi $c_1 \equiv c_2$
- $\sigma'(c_1, c_2) = 0$ khi hoặc $c_1 \equiv \perp$ hoặc $c_2 \equiv \perp$.

Ví dụ dưới đây sẽ cho thấy tính chất ở Mệnh đề 1 của độ tương tự σ'_O cũng có trong độ tương tự σ'_O .

Ví dụ 3. Xét ontology như ở Ví dụ 1. Chúng ta tính độ tương tự ngữ nghĩa σ' cho các cặp khái niệm $(Lion, Cat)$ và $(Lion, Animal)$.

$$\sigma'(Lion, Cat) =$$

$$\frac{2 \cdot \mu'(Animal, T)}{\mu'(Lion, Animal) + \mu'(Cat, Animal) + 2 \cdot \mu'(Animal, T)} = \frac{2 \cdot 2}{1 + 1 + 2 \cdot 2} \approx 0.66$$

$$\sigma'(Sheep, Cat) =$$

$$\frac{2 \cdot \mu'(Animal, T)}{\mu'(Sheep, Animal) + \mu'(Cat, Animal) + 2 \cdot \mu'(Animal, T)} = \frac{2 \cdot 2}{2 + 1 + 2 \cdot 2} \approx 0.57$$

Như vậy, $\sigma'(Lion, Cat) > \sigma'(Sheep, Cat)$ – hay, *Lion*

giống với *Cat* hơn khi so sánh *Sheep* và *Cat*. Điều phản ánh đúng hình ảnh trực quan của cây phân cấp khái niệm: Hai khái niệm càng cách xa khái niệm cha chung nhỏ nhất thì có độ giống nhau càng thấp. Nói cách khác, độ tương tự σ' vẫn giữ đúng ưu điểm của phương pháp tính Wu-Palmer.

3.1. Khảo sát độ tương tự σ' của hai khái niệm khi có khái niệm cha chung nhỏ nhất là khái niệm đỉnh

Xét hai khái niệm $c_1, c_2 \in \mathcal{C}_O$ sao cho $LCP_O(c_1, c_2) = \{T\}$. Khi đó $\sigma'(c_1, c_2)$ được xác định như sau:

$$\sigma'_O(c_1, c_2) = \frac{2 \cdot \mu'_O(T, T)}{\mu'_O(c_1, T) + \mu'_O(c_2, T) + 2 \cdot \mu'_O(T, T)} = \frac{2}{\mu_O(c_1, T) + \mu_O(c_2, T) + 4}$$

Điều này chứng tỏ: Khi $LCP_O(c_1, c_2) = \{T\}$ thì $\sigma'(c_1, c_2) > 0$ chứ không bị triệt tiêu như phương pháp của Wu-Palmer. Hơn nữa, $\sigma'(c_1, c_2)$ cũng bảo đảm rằng, khi hai khái niệm c_1, c_2 càng cách xa khái niệm đỉnh T (và càng cách xa nhau – do hai khái niệm thuộc về hai nhánh của khái niệm đỉnh T) thì độ tương tự giữa chúng càng giảm. Điều này cũng phản ánh đúng tính chất trực quan của cây phân cấp khái niệm.

Ví dụ 4. Xét lại ontology ở Ví dụ 1. Ta có:

$$LCP_O(\text{BodyPart}, \text{Cat}) =$$

- $LCP_O(\text{BodyPart}, \text{Animal}) = \{T\}$

$$\sigma'(\text{BodyPart}, \text{Cat}) = \frac{2 \cdot \mu'(T, T)}{\mu'(\text{BodyPart}, T) + \mu'(\text{Cat}, T) + 2 \cdot \mu'(T, T)} = \frac{2 \cdot 1}{2 + 1 + 2 \cdot 1} = \frac{2}{5}$$

$$\sigma'(\text{BodyPart}, \text{Animal}) = \frac{2 \cdot \mu'(T, T)}{\mu'(\text{BodyPart}, T) + \mu'(\text{Animal}, T) + 2 \cdot \mu'(T, T)} = \frac{2 \cdot 1}{2 + 2 + 2 \cdot 1} = \frac{2}{6}$$

Như vậy, $\mu'(\text{BodyPart}, \text{Cat}) < \mu'(\text{BodyPart}, \text{Animal})$ – hay, *BodyPart* giống với *Animal* hơn khi so sánh *BodyPart* với *Cat*. Điều này cũng phản ánh đúng hình ảnh trực quan của cây phân cấp khái niệm ở Hình 1.

3.2. Vấn đề tính độ tương tự ngữ nghĩa giữa hai hai biểu thức khái niệm

Với O là một ontology, chúng ta mở rộng khái niệm độ tương tự giữa hai khái niệm trong ontology thành “độ tương tự giữa hai biểu thức khái niệm theo ontology O ” như sau: Độ tương tự ngữ nghĩa giữa hai biểu thức khái niệm

CE_1, CE_2 theo ontology O được hiểu là độ tương tự của hai biểu thức khái niệm này khi đặt trên cây phân cấp khái niệm của ontology O . Ontology O khi đó được gọi là ontology tham chiếu.

Độ tương tự ngữ nghĩa giữa hai khái niệm CE_1, CE_2 theo ontology O được xác định theo công thức ở Định nghĩa 4 như sau:

$$\sigma'_O(CE_1, CE_2) = \min \left\{ \frac{2 \cdot \mu'_O(CE_0, T)}{\mu'_O(CE_1, CE_0) + \mu'_O(CE_2, CE_0) + 2 \cdot \mu'_O(CE_0, T)} \mid CE_0 \in LCP_O(CE_1, CE_2) \right\}$$

Chúng ta có thể tính số cung nối giữa hai biểu thức khái niệm bất kỳ CE_1 và CE_2 dựa theo số cung nối giữa hai khái niệm có tên của ontology O theo 3 trường hợp như sau:

- Trường hợp 1) $\exists c_1 \in \mathcal{C}_O: c_1 \equiv CE_1$. Khi đó: $\mu'(CE_1, CE_2) = \min \{ \mu'_O(c_1, c) \mid c \in DC_O(CE_2) \}$
- Trường hợp 2) $\exists c_2 \in \mathcal{C}_O: c_2 \equiv CE_2$. Khi đó: $\mu'_O(CE_1, CE_2) = \min \{ \mu'_O(c, c_2) \mid c \in DP_O(CE_1) \}$
- Trường hợp 3) Không có khái niệm có tên trên O tương đương với CE_1 hoặc CE_2 . Khi đó:

$$\mu'_O(CE_1, CE_2) = \min \{ 2 + \mu'_O(c, c') \mid c \in DP_O(CE_1), c' \in DC_O(CE_2) \}$$

Như vậy, bằng cách tính sẵn các thông số μ' giữa các cặp khái niệm của ontology O (có thể dùng thuật toán tìm đường đi ngắn nhất giữa các cặp đỉnh như Floyd [1] chẳng hạn), chúng ta có thể tính nhanh độ tương tự ngữ nghĩa σ' theo ontology tham chiếu O của cặp biểu thức khái niệm bất kỳ.

IV. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Bài báo đã đề xuất cách tính số cung nối μ' giữa hai khái niệm trên cây phân cấp khái niệm của ontology, qua đó đưa ra cải tiến cho phương pháp Wu-Palmer để tính độ tương tự ngữ nghĩa σ' giữa hai khái niệm. Độ tương tự này vẫn giữ nguyên ưu điểm của phương pháp gốc, nhưng cho phép định lượng được mức độ giống nhau của các cặp khái niệm nhận T làm khái niệm cha chung nhỏ nhất. Bài báo cũng đưa ra phương pháp hiệu quả để tính độ tương tự ngữ nghĩa theo ontology tham chiếu của hai biểu thức khái niệm bất kỳ.

Trong tương lai, chúng tôi sẽ phân tích và áp dụng các đề xuất của bài báo này cho các phương pháp đánh giá độ tương tự ngữ nghĩa khác (ngoài Wu-Palmer) theo hai hướng: đánh giá độ tương tự ngữ nghĩa giữa hai khái niệm trong ontology và đánh giá độ tương tự ngữ nghĩa giữa hai biểu thức khái niệm bất kỳ.

REFERENCES

- [1] R. Floyd (1962). Algorithm 97: shortest path, *Communications of the ACM*, Vol. 5, Issue 6, pp. 345.
- [2] T. Slimani (2013). Description and Evaluation of Semantic Similarity Measures Approaches. *International Journal of Computer Applications*, Vol. 80, Issue 10, pp. 25-33.
- [3] T. V. Nguyen, H. H. Hoang (2016). A Consensus-Based Method for Solving Concept-Level Conflict in Ontology Integration, *Trans. Computational Collective Intelligence*, Vol. LCNS 8733, Issue XXII, pp. 414-423.

- [4] T. Wu, Z. Palmer (1994), Verb Semantics and Lexical Selection, *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, pp. 133-138.
- [5] Zhisheng Huang, Frank van Harmelen (2008). Using semantic distances for reasoning with inconsistent ontologies, *The Semantic Web - ISWC 2008*, p. 454-459.



Hoàng Hữu Hạnh sinh ngày 13/04/1974 tại Huế. Năm 1996, ông tốt nghiệp Cử nhân ngành Toán-Tin học tại Trường Đại học Sư phạm Huế, Thạc sĩ khoa học tại Trường Đại học Bách khoa Hà Nội. Năm 2007, ông nhận học vị Tiến sĩ chuyên ngành Hệ thống thông tin tại Trường Đại học Công nghệ Vienna, Cộng hòa Áo. Năm 2012, ông nhận Chức

danh Phó giáo sư tại Việt Nam. Từ năm 1996-2018 đến nay, ông là Giảng viên Khoa Công nghệ Thông tin, Trường ĐH Khoa học Huế; đồng thời đảm trách các nhiệm vụ quản lý tại Đại học Huế từ 2008 đến 2018. Từ tháng 10 năm 2018 ông là Giảng viên cao cấp tại Học viện Công nghệ Bưu chính Viễn thông (PTIT). Hiện nay là Giám đốc Trung tâm Đào tạo Quốc tế của PTIT.

Lĩnh vực nghiên cứu: Biểu diễn tri thức, Web ngữ nghĩa, Linked Data, Ontology, Logic mô tả, Công nghệ phần mềm, Công nghệ dữ liệu, Quản lý quy trình nghiệp vụ.



Nguyễn Văn Trung sinh ngày 25/10/1981 tại Thừa Thiên Huế. Năm 2003 ông tốt nghiệp cử nhân chuyên ngành Tin học tại trường Đại học Khoa học Huế. Năm 2018 ông nhận bằng Tiến sĩ chuyên ngành Khoa học máy tính tại trường Đại học Huế. Từ năm 2004 đến nay ông giảng dạy và nghiên cứu khoa học tại Khoa Công nghệ Thông tin, trường Đại học Khoa học Huế.

Lĩnh vực nghiên cứu: Các hệ thống thông tin, Quản lý và biểu diễn tri thức, Công nghệ phần mềm.