Ontologies – Future Perspectives

Semantic Web

F. Abel, N. Henze, and D. Krause

IVS Semantic Web Group

December 12, 2008

1 Exercise 1: FolkRank

Given are the following tag assignments:

```
(u1, t1, r1)
(u1, t2, r1)
(u1, t2, r2)
(u2, t2, r2)
```

1. Compute the FolkRank for the folksonomy induced by the given tag assignments (only perform the first Iteration, $d = 0.5$). The Vector $\vec{w}$ should be initialized as follows:

$$\vec{p} = \begin{pmatrix} \text{ranking score of } u_1 \\ \text{ranking score of } u_2 \\ \text{ranking score of } t_1 \\ \text{ranking score of } t_2 \\ \text{ranking score of } r_1 \\ \text{ranking score of } r_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}$$

The preference vector is given as follows:

$$\vec{p} = \begin{pmatrix} \text{preference in } u_1 \\ \text{preference in } u_2 \\ \text{preference in } t_1 \\ \text{preference in } t_2 \\ \text{preference in } r_1 \\ \text{preference in } r_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 21 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Note that $\vec{p}$ has to fulfill the condition $||\vec{w}||_1 = ||\vec{p}||_1$.

1.1 Solution

From the given tag assignments we construct a graph, which is then modeled within the adjacency matrix A. Therefore, we construct A by going through the given tag assignments step by step.

Construction of Adjacency Matrix A:
The dimensions of A have to correspond to the dimensions of the preference

vector $\vec{p}$ and the vector $\vec{w}$, i.e.:

	u_1	u_2	t_1	t_2	r_1	r_2
u_1						
u_2						
t_1						
t_2						
r_1						
r_2						

Step 1 ((u_1, t_1, r_1)):

	u_1	u_2	t_1	t_2	r_1	r_2
u_1			1		1	
u_2						
t_1	1				1	
t_2						
r_1	1		1			
r_2						

Step 2 ((u_1, t_2, r_1)):

	u_1	u_2	t_1	t_2	r_1	r_2
u_1			1	1	2	
u_2						
t_1	1				1	
t_2	1				1	
r_1	2		1	1		
r_2						

Step 3 ((u_1, t_2, r_2)):

	u_1	u_2	t_1	t_2	r_1	r_2
u_1			1	2	2	1
u_2						
t_1	1				1	
t_2	2				1	1
r_1	2		1	1		
r_2	1			1		

Step 4 ((u_2, t_2, r_2)):

	u_1	u_2	t_1	t_2	r_1	r_2
u_1			1	2	2	1
u_2				1		1
t_1	1				1	
t_2	2	1			1	2
r_1	2		1	1		
r_2	1	1		2		

Step 5 – Build the 1-norm for each row:

	u_1	u_2	t_1	t_2	r_1	r_2
u_1			$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{1}{6}$
u_2				0.5		0.5
t_1	0.5				0.5	
t_2	$\frac{2}{6}$	$\frac{1}{6}$			$\frac{1}{6}$	$\frac{2}{6}$
r_1	0.5		0.25	0.25		
r_2	0.25	0.25		0.5		

Compute $\vec{w}_0$
For the computation of $\vec{w}_0$ we have to set $d = 1$ so that there is no influence of the preference vector ($\vec{w}_0 \leftarrow A \cdot \vec{w}_0$).

Iteration 1:

$$\vec{w}_{0,1} = A \cdot \vec{w}_{0,0} = \begin{pmatrix} 0 & 0 & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0 & 0.5 & 0 \\ \frac{2}{6} & \frac{1}{6} & 0 & 0 & \frac{1}{6} & \frac{2}{6} \\ 0.5 & 0 & 0.25 & 0.25 & 0 & 0 \\ 0.25 & 0.25 & 0 & 0.5 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 4.5 \\ 5 \\ 3 \\ 3.5 \\ 2.25 \\ 2.75 \end{pmatrix}$$

Compute $\vec{w}_1$
For the computation of $\vec{w}_1$ we work with the given value $d = 0.5$ and the given preference vector ($\vec{w}_1 \leftarrow dA \cdot \vec{w}_1 + (1-d)\vec{p}$).

Iteration 1:

$$\vec{w}_{1,1} = dA \cdot \vec{w}_{1,0} + (1-d)\vec{p} = 0.5 \cdot \begin{pmatrix} 0 & 0 & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0.5 & 0 & 0 & 0 & 0.5 & 0 \\ \frac{2}{6} & \frac{1}{6} & 0 & 0 & \frac{1}{6} & \frac{2}{6} \\ 0.5 & 0 & 0.25 & 0.25 & 0 & 0 \\ 0.25 & 0.25 & 0 & 0.5 & 0 & 0 \end{pmatrix} \cdot$$

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix} + 0.5 \cdot \begin{pmatrix} 0 \\ 0 \\ 21 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2.25 \\ 2.5 \\ 12 \\ 1.75 \\ 1.125 \\ 1.375 \end{pmatrix}$$

Compute $\vec{w}$

$$\vec{w} = \vec{w}_1 - \vec{w}_0 = \begin{pmatrix} 2.25 \\ 2.5 \\ 12 \\ 1.75 \\ 1.125 \\ 1.375 \end{pmatrix} - \begin{pmatrix} 4.5 \\ 5 \\ 3 \\ 3.5 \\ 2.25 \\ 2.75 \end{pmatrix} = \begin{pmatrix} -2.25 \\ -2.5 \\ 9 \\ -1.75 \\ -1.125 \\ -1.375 \end{pmatrix}$$

2 Exercise 2: Personomy, Association Rules

Given are the following tag assignments:

```
(u1, jazz, r1)
(u1, trumpet, r1)
(u2, music, r1)
(u2, jazz, r2)
(u3, chet-baker, r2)
(u3, music, r2)
(u1, classic, r3)
(u2, music, r3)
```

1. Write down the *Personomies* for each user.
2. Detect two "good" association rules.

2.1 Solution – Personomies

Personomy of user u_1: $P_{u1} := (T_{u1}, R_{u1}, Y_{u1})$, where:

- $T_{u1} = \{\text{jazz, trumpet, classic}\}$
- $R_{u1} = \{r_1, r_3\}$
- $Y_{u1} = \{(\text{jazz}, r_1), (\text{trumpet}, r_1), (\text{classic}, r_3)\}$

Personomy of user u_2: $P_{u2} := (T_{u2}, R_{u2}, Y_{u2})$, where:

- $T_{u2} = \{\text{music, jazz}\}$
- $R_{u2} = \{r_1, r_2, r_3\}$
- $Y_{u2} = \{(\text{music}, r_1), (\text{jazz}, r_2), (\text{music}, r_3)\}$

Personomy of user u_3: $P_{u3} := (T_{u3}, R_{u3}, Y_{u3})$, where:

- $T_{u3} = \{\text{chet-baker, music}\}$
- $R_{u3} = \{r_2\}$
- $Y_{u3} = \{(\text{chet-baker}, r_2), (\text{music}, r_2)\}$

2.2 Solution – Association Rules

The set of items is equal to the set of tags: $I = T = \{$jazz, trumpet, music, chet-baker, classic$\}$

Transactions: $D = \{D_{r1}, D_{r2}, D_{r3}\}$, where $D_{r1} = \{$jazz, trumpet, music$\}$, $D_{r2} = \{$jazz, chet-baker, music$\}$, and $D_{r1} = \{$classic, music$\}$.

Examples for association rules:

1. music $\Longrightarrow$ jazz

support(music $\Longrightarrow$ jazz) $= \frac{2}{3}$, confidence(music $\Longrightarrow$ jazz) $= \frac{2}{3}$

2. jazz $\Longrightarrow$ music

support(jazz $\Longrightarrow$ music) $= \frac{2}{3}$, confidence(jazz $\Longrightarrow$ music) $= \frac{2}{2} = 1$

3. trumpet $\Longrightarrow$ classic

support(trumpet $\Longrightarrow$ classic) $= \frac{0}{3} = 0$, confidence(trumpet $\Longrightarrow$ classic) $= \frac{0}{1} = 0$

4. classic $\Longrightarrow$ music

support(classic $\Longrightarrow$ music) $= \frac{1}{3}$, confidence(classic $\Longrightarrow$ music) $= \frac{1}{1}$

Discussion:

If we compare rule 1 and 2 then we can clearly say that *"jazz $\Longrightarrow$ music"* (rule 2) is a better rule than *"music $\Longrightarrow$ jazz"* (rule 1). Both rules have the same support level. However, the confidence of rule 2 is higher than the one of rule 1. In particular, *"confidence(jazz $\Longrightarrow$ music) = 1"* means that whenever "jazz" occurs then "music" occurs as well.

The hypothesis of the ontology learning approach, which is based on association rules, is the following: An association rule *"$tag_x \Longrightarrow tag_y$"* learned from a folksonomy can be interpreted as *"tag_x is a sub-concept of tag_y"*. For our example, we could thus say that *"jazz is a sub-concept of music"*, which makes sense.

Assumption of the ontology learning approach, which is based on association rules: More general tags occur more often than specialized tags. Hence, the probability that "*confidence(super-tag $\Longrightarrow$ sub-tag) $\leq$ confidence(sub-tag $\Longrightarrow$ super-tag)*" is higher than "*confidence(sub-tag $\Longrightarrow$ super-tag) $\leq$ confidence(super-tag $\Longrightarrow$ sub-tag)*".