

저널 논문 투고 및 심사 시스템에서 심사자 추천 알고리즘

정용진, 김용환, 김찬명, 한연희+
한국기술교육대학교 컴퓨터공학부/첨단기술연구소
{jyjin989, cherish, cmdr, yhhan}@koreatech.ac.kr

A Reviewer Recommendation Algorithm in Journal Submission and Review Systems

Yong-Jin Jeong, Yong-hwan Kim, Chan-Myung Kim and Youn-Hee Han+
School of Computer Science and Engineering
/Advanced Technology Research Center,
Korea University of Technology and Education

요 약

저널 논문 투고 및 심사시스템에서의 논문 제출은 상시 이루어진다는 특성 때문에 논문이 제출된 시점에 적절한 심사자들을 찾아 배정하기란 쉽지 않은 문제이다. 본 논문에서는 이러한 문제를 해결하기 위하여 제출된 논문에 적절한 심사자들을 추천해주는 알고리즘을 제시하고자 한다. 심사자 추천 알고리즘에서는 해당 논문의 전문가를 심사자로서 추천하기 위하여 제출된 논문들의 키워드(Keyword)와 심사자들의 전문지식태그(Expertise Tag) 정보를 활용한다. 또한, 심사자들의 기존의 심사 정보를 토대로 심사활동지수를 평가하여 이를 심사자 추천에 활용하고자 한다. 제안하는 알고리즘을 검증하기 위하여 본 논문에서는 실제 저널 논문투고시스템에 추천 알고리즘을 적용해보고 이의 결과를 제시한다.

1. 서론

본 논문은 저널 논문투고시스템에서 제출된 논문들에 대하여 적절한 심사자들을 추천하는 방안에 대하여 제시하고자 한다. 이는 논문의 제출이 상시 이루어지는 저널 논문투고시스템의 관리 및 처리과정을 더욱 용이하도록 도움을 주는데 중요한 역할을 하게 된다 [1].

기존의 많은 심사자 추천 시스템들에서는 최대한 해당 논문의 전문가를 심사자로 추천하는데 주안점을 두고 있다[2, 3, 4]. 그러나 기존의 추천 시스템은 추천의 우선순위가 나타나지 않으며 심사자들의 참여빈도를 고려하지 않고 있다. 이에 따라, 소수의 심사자들에게 논문 심사가 집중되어 제한 기간 내에 심사가 원활하지 이뤄지지 못하거나 특정 연구자들에게는 심사기회가 잘 주어지지 않는 상황이 발생되기도 한다. 또한, 기존의 심사자 추천 알고리즘의 경우 알고리즘을 수행하는데 소요되는 시간이 적지 않다는 단점이 있다.

본 논문에서는 논문의 키워드와 심사자들의 전문지식태그 정보를 기반으로 전문 분야의 유사성을 측정하여 보다 논문의 주제와 유사하고 전문적인 심사자들을 추천한다. 또한, 심사자들의 최근 심사 참여에 대한 수치를 나타낸 심사활동지수를 활용하여 심사자들의 심사 참여 정도를 최대한 균등하게 하고자한다.

2장에서는 논문에서 해결하고자 하는 문제의 정의와 이에 대한 해결 방안을 제시하고, 3장에서는 실제 한국정보

처리학회 저널 논문 투고 및 심사시스템[5]에 제안 알고리즘을 적용한 결과를 제시하며, 마지막으로 4장에서 본 논문에 대한 결론을 짓는다.

2. 논문 심사자 추천 문제 및 제안 알고리즘

본 논문에서 해결하고자 하는 논문 심사자 추천 문제는 다음과 같다.

[논문 심사자 추천 문제] 저널 논문 투고 시스템에서 제출된 논문들의 집합 A 와 심사자들의 집합 R 이 존재할 때, 임의의 논문 $a \in A$ 마다 적절한 심사자 $r \in R$ 들을 찾아 추천하는 문제이다. 이의 문제는 다음과 같은 제한사항들을 만족해야 한다.

- 1) θ_1 - 제출논문은 시스템에 의해 규정된 서로 다른 l 명의 심사자들에게 할당 되어야 한다.
- 2) θ_2 - 각 심사자 r 은 최대 l 개의 논문을 심사 할 수 있다.
- 3) θ_3 - 논문 a 에 대하여 심사가 가능한 자격이 충족된 심사자 r 을 추천 할 수 있다.
- 4) θ_4 - 각 심사자 r 은 최대 $m_r(l_r - \text{현재 심사 중인 수})$ 번 추천 될 수 있다.

위의 문제를 해결하기 위하여 본 논문에서는 논문과 심사자 사이의 전문분야 유사성과 심사자의 심사 참여빈도를 계산하여 해당 논문에 적합한 심사자들을 추천한다 [2].

다음 식 (1)은 임의의 논문 a 의 키워드와 임의의 심사

+ 한연희: 교신저자

자 의 전문지식태그가 얼마나 많은 유사한 주제를 가지는지 정도를 측정하기 위한 공통주제 평가함수 를 나타낸다.

$$U(a,r) = \frac{|K_1(r) \cap K_2(a)|}{|K(a)|} \quad (1)$$

- $K_1(a)$ - 논문 의 키워드의 집합
 - $K_2(r)$ - 심사자 r 의 전문지식태그의 집합
- 식 (1)에서 키워드와 전문지식태그의 주제가 유사한지 판단하는 기준은 다음과 같다.

- 1) 주제어 전체의 80%이상 일치 할 경우에 공통주제로 간주한다. ex) social network, social networks
- 2) 주제어가 다수의 단어로 이루어진 경우, 과반수의 단어들의 조합을 추출하여 해당 단어조합의 80%이상 일치 할 경우에 공통주제로 간주한다.

식 (2)에서는 논문 a 의 심사에 심사자 r 을 추천하는 판단에 있어 기존의 논문 a 의 심사자 그룹 R_a 에 새로운 심사자 r 이 참여하여 구성되는 새로운 심사자 그룹 R_a 가 지닌 전문지식 집합과 논문 a 의 키워드 집합 간의 유사성을 비교하고, 그룹 R_a 의 심사자들이 대체적으로 비슷한 수준의 주제유사성을 가지고 있는지 측정한다. 이를 그룹 공통주제 평가함수 라 하며 다음과 같이 정의한다.

$$V(a,r) = \frac{|K_1(a) \cap \bigcup_{s \in R_a \cup \{r\}} K_2(s)|}{|K(a)|} - e \quad (2)$$

위 식에서 e 는 다음과 같이 정의된다.

$$e = \frac{\sum_{s \in R_a \cup \{r\}} |K_1(a) \cap K_2(s)|}{|R_a \cup \{r\}|}$$

- R_a - 논문 a 를 심사 중인 심사자들의 집합

식 (3)은 공통주제 평가함수와 그룹공통주제 평가함수를 기반으로 논문과 심사자 사이의 전문분야 유사성을 측정하는 심사 추천 평가함수를 나타낸다.

$$(a,r) = \gamma \cdot U(a,r) + (1-\gamma) \cdot V(a,r) \quad (3)$$

- γ - 심사 추천 평가함수 내부 가중치 파라미터

식 (4)는 최근 특정 기간 동안의 심사 참여빈도를 계산한 심사 활동 지수를 나타낸다. 임의의 논문 a 에 대하여 임의의 두 심사자가 비슷한 전문분야 유사성을 지닐 경우,

심사 참여빈도가 낮은 심사자를 추천하여 심사자들 간의 심사기회를 균등하게 부여하도록 한다.

$$FR_r = \frac{A_r}{T} \quad (4)$$

- T - 최근 특정 기간
- A_r - T 기간 동안 심사자 r 이 심사에 참여한 횟수.

다음은 논문 심사자 추천 문제를 해결하기 위해 위의 식 (3)과 (4), 다시 말하면 전문분야 유사성과 심사자의 심사참여빈도를 고려한 심사자 추천 알고리즘이다.

Algorithm 1

1. **Input:** A, R
 2. **output:** *recommended_reviewer_set*
 3. **for** $a \in A$ **do**
 4. **for** $r \in R$ **do**
 5. **if** $(f(a,r) > \epsilon$ **and** $b_{ar} = true)$ **then**
 6. *candidate_reviewer_set.add(r);*
 7. **end if**
 8. **end for**
 9. **for** $i \in [1 \dots (k \cdot v)]$ **do**
 10. $et r \in \text{argmin} \in R \{r \in \text{candidate_reviewer_set}\};$
 11. *recommended_reviewer_set.add(r);*
 12. **end for**
 13. **end for**
-

3-8번째 줄에서는 심사가 요구되는 제출된 논문들의 집합 A 에 대해, 주제유사성을 기반으로 하여 심사 추천의 적합성을 나타내는 식 (3)을 이용하여 유사성 정도가 임계값 ϵ 보다 높은 심사자들을 추천 후보 집합에 넣는다. 이때, b_{ar} 는 논문 심사자 추천 문제의 제한사항을 만족하는지에 대한 여부를 나타내며 θ_3 의 경우, 심사자 r 가 논문 a 에 대해 이미 심사 중인지, 심사자 r 이 논문 a 의 저자와 공동으로 논문을 작성한 적이 있는지 등의 심사 자격여부를 따진다.

9-12번째 줄에서는 식 (4)를 이용하여 3-8번째 줄에서 선정된 추천 후보들 가운데 최근 심사 참여빈도가 적은 심사자를 추천한다. 이 때 최대 추천 가능한 심사자들의 수는 제한사항 θ_1 에서 규정된 k 에 v 배수로 제한한다. 한편, 제안 알고리즘에서 심사 활동 지수를 고려하여 추천된 심사자들의 우선순위는 $f(a,r)$ 에 의하여 결정한다.

3. 심사자 추천 알고리즘 적용 결과

본 장에서는 제안하는 논문 심사자 추천 알고리즘을 검

증하기 위하여 실제 한국 정보처리학회 저널 논문 투고 및 심사시스템 [5]에 해당 알고리즘을 적용하여 도출된 결과 중 일부를 제시한다. 해당 실험에서 사용된 파라미터 값은 각각 $\alpha = 3, l = 3, v = 2, \gamma = 0.75, \epsilon = 0.3$ 이다.

<표 1> 추천 결과 목록

심사 논문	(추천순위) 추천심사자	키워드 - 전문지식태그
683		Core-set Gaussian Mixture Models Minimal Enclosing Ball Support Vector Machines Universal Background Model
	(1) 515	Core-set Gaussian Background Mixture Models Svm (Support Vector Machine) ⋮
	(2) 1197	Gaussian Background Mixture Models Svm (Support Vector Machine) ⋮
	(3) 323	Gaussian Background Mixture Models Support Vector Machine ⋮
	(4) 1258	Gaussian Mixture Model (Gmm) Support Vector Machines ⋮
	(5) 326	Gaussian Mixture Model (Gmm) Support Vector Machines ⋮
784		Cloud Computing CloudAnalyst Data center, Virtualisation load balancer Virtual Machine
	(1) 423	Cloud Computing Cloudanalyst VM Load Balancer Virtual Machine ⋮
	(2) 221	Cloudanalyst VM Load Balancer Virtual Machine ⋮
	(3) 258	Cloudanalyst VM Load Balancer Virtual Machine ⋮
	(4) 230	Cloudanalyst VM Load Balancer Virtual Machine ⋮
799		Ant-colony Cognitive behavior Pheromone Psychology of programming Social network and collective intelligence Software Development Life Cycle
	(1) 215	Social Networking Sites Agile Software Development Software Development Methodology ⋮

(2) 569	Agile Software Development Software Development Methodology ⋮
(3) 1109	Agile Software Development Software Development Methodology ⋮
(4) 298	Agile Software Development Software Development Methodology ⋮
(5) 2348	Agile Software Development Software Development Methodology ⋮
(6) 1033	Agile Software Development Software Development Methodology ⋮

<표 1>은 실제 저널 논문투고시스템에 제안 알고리즘 적용 후 도출된 심사자 추천 결과목록 중 임의의 3가지 예시의 추천 결과를 보여준다. 추천 결과를 살펴보면 논문의 키워드와 유사한 전문지식태그를 보유하고 있는 심사자들을 추천하고 있으며 더 많은 공통 주제를 가진 심사자에게 높은 추천 순위를 부여함을 볼 수 있다. 또한, 해당분야에 있어 최근 심사참여빈도를 고려하여 참여빈도가 적은 심사자들을 추천함으로써 균등한 심사참여기회를 제공한다.

4. 결론

본 논문은 저널 논문투고시스템에서 제출된 논문에 대한 심사자들을 추천하는 방법을 제시했다. 주제유사성 비교를 기반으로 하여 해당 논문의 분야에 보다 더 전문적인 심사자를 추천하도록 한다. 또한 기존의 추천 시스템과 달리 우선순위를 나타내며, 또한 심사참여빈도를 고려하여 심사자들의 심사기회가 균등하게 이루어지도록 하였다. 이러한 방안이 논문투고시스템의 관리 및 처리과정을 용이하도록 하는데 도움이 될 것이라 기대한다.

참고문헌

- [1] Donald C. Conry "Recommender Systems for the Conference Paper Assignment Problem" the ACM Conference on Recommender Systems 2009.
- [2] Tomasz Kolasa and Dariusz Krol "A Survey of Algorithms for Paper-reviewer Assignment Problem" IETE Technical Review (Medknow Publications & media Pvt. Ltd.) 2011.
- [3] Xu Yun-hong, GUO Xi-tong, XU Liang, CHEN Yu, ZHUANG Yong-yao "Research Analytics for Reviewer Recommendation" 2012 International Conference on Management Science & Engineering (19th).
- [4] 이정연, 이재윤, 정한민, 강인수, 신숙경 "확률적 온톨로지와 연구자 네트워크를 이용한 심사자 자동 추천에 관한 연구" 정보관리학회지 Vol. 24 No. 3 2007.
- [5] <http://www.manuscriptlink.com/journals/jips>