



## 중장년 직업추천 알고리즘 개발을 위한 중장년층 직종 선호 및 취업 특성 분석

Job Preference and Selection Analysis of Senior People for Developing Senior Job Recommendation Algorithm

---

저자 (Authors)	김성중, 김성찬, 장진철, 진효진, 이문용 SeongJung Kim, Seongchan Kim, Jincheul Jang, Hyojin Chin, Mun Y. Yi
출처 (Source)	<a href="#">한국정보과학회 학술발표논문집</a> , 2016.6, 331-333 (3 pages)
발행처 (Publisher)	<a href="#">한국정보과학회</a> KOREA INFORMATION SCIENCE SOCIETY
URL	<a href="http://www.dbpia.co.kr/Article/NODE07017490">http://www.dbpia.co.kr/Article/NODE07017490</a>
APA Style	김성중, 김성찬, 장진철, 진효진, 이문용 (2016). 중장년 직업추천 알고리즘 개발을 위한 중장년층 직종 선호 및 취업 특성 분석. 한국정보과학회 학술발표논문집, 331-333.
이용정보 (Accessed)	KAIST 143.248.91.*** 2018/03/20 16:13 (KST)

---

### 저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독 계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

### Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

# 중장년 직업추천 알고리즘 개발을 위한 중장년층 직종 선호 및 취업 특성 분석

김성중<sup>01</sup> 김성찬<sup>1</sup> 장진철<sup>1</sup> 진효진<sup>1</sup> 이문용<sup>1</sup>

<sup>1</sup>한국과학기술원 지식서비스공학대학원

myuse0120@kaist.ac.kr, sckim@kaist.ac.kr, jcjang@kaist.ac.kr, tesschin@kaist.ac.kr, munyi@kaist.ac.kr

## Job Preference and Selection Analysis of Senior People for Developing Senior Job Recommendation Algorithm

SeongJung Kim<sup>01</sup> Seongchan Kim<sup>1</sup> Jincheul Jang<sup>1</sup> Hyojin Chin<sup>1</sup> Mun Y. Yi<sup>1</sup>

<sup>1</sup>Graduate School of Knowledge Service Engineering, KAIST

### 요약

전체 인구 중 고령층의 비중이 꾸준히 높아지며, 인구고령화는 한국 사회에서 점점 더 중요한 이슈로 부각되고 있다. 인구고령화는 사회 전체적으로 노동력 재편을 요구하는 중요한 이슈이다. 조기 은퇴 및 기대수명의 증가로 인해 제 2의 직업을 찾는 중장년층에게 적절한 일자리 정보를 제공하는 일은 사회 전체의 생산성 향상 및 경제 성장과 직결되어 있으며, 사회 안정화에 커다란 영향을 갖는다. 최근 활발하게 진행되고 있는 개인화 추천 연구는 아직까지 중장년층에 특화된 직업 추천 알고리즘을 제시하지 못하고 있다. 이러한 상황에서 본 연구에서는 중장년층을 대상으로 한 직업 추천 알고리즘 개발을 목적으로 3년간 축적한 182,516개의 워크넷 실제 구직자 데이터를 분석하였다. 이러한 분석을 바탕으로 중장년층의 선호 직종과 취업특성을 찾아내고 향후 중장년층 직업 추천 알고리즘에 반영하고자 한다.

### 1. 서론

저출산과 함께 증가된 기대수명으로 인해 전체 인구 중 고령 인구 비중이 꾸준히 상승하면서 인구 고령화는 우리나라의 중요한 사회적 이슈로 부각되어 왔다. 통계청에 따르면, 2000년을 기점으로 우리나라 65세 이상 고령인구가 전체 인구의 7%를 넘어섰으며, 2017년에는 두 배인 14%가 될 것으로 예측하고 있다. 특히 2015년 기준 이미 전남 지역은 고령인구 비율이 20%를 넘어섰으며 전북, 경북 순으로 20%에 육박한다고 조사되었다. 또한, 고령자의 최장근속기간은 2006년 평균 17년 3.1개월이었던 것에 비해 2015년에는 14년 9.4개월로 지속적으로 감소하는 추세이며, 2010년대에 베이비붐 세대(1955-1963년생)의 정년시기에 도래하면서 이들의 대규모 은퇴로 인해 새로운 직업을 찾기 원하는 중장년층의 수가 급속도로 증가하였다 [1]. 이에 따라, 중장년층 구직자에게 새로운 직업을 효율적으로 제공해줌으로써 경제 활동인구, 노동 생산성 등을 향상시키며, 경제 성장률 둔화 등의 문제점을 해결해야 할 필요성이 높아지고 있다.

이러한 상황과 더불어 최근 스마트폰의 대중화와 데이터 처리 기술, 사용자 편의 중심 인터페이스의 발달은 IT기술이 중장년층에게 새로운 일자리를 제공하는 유용한 수단이 될 수 있음을 보여주고 있다. 한 예로, 정부에서 운영하는 워크넷<sup>1)</sup> 웹사이트의 통계 결과를 보면, 2000년 3월 한 달 동안 65세 이상 신규 구직자 수가 3,882명인 것에 비해 2015년 3월에는 30,236명인 것으로 조사되었다 [2]. 또한, 잡코리아, 알바천국 등 민간 기업에서도 구인 구직 서비스를 운영하고 있으며, 과거 직업 상담 시설의 상담을 통한 구인 구직을 넘어서 웹을 활용한 새로운 형태의 직업 알선 서비스가 자리 잡고 있다.

또한, 90년대 이후 개인화 추천 알고리즘이 지속적으로 연구되어 왔다. 이러한 연구는 직업 추천 분야에도 적용되었으며, 기존 추천 알고리즘으로 해결하지 못하는 직업 추천 분야에 특화된 알고리즘을 개발하기 위한 연구가 진행 중에 있다. 직업 추천의 경우 기존 추천 알고리즘과 다르게 한 사용자가 선택하는 직업의 수가 상대적으로 적은 문제가 있으며, 또한 구직자와 구인자 모두에게 만족스러운 추천 결과여야하기 때문에 기존 추천 알고리즘과 다른 성격을 띤다. 기존의 직업 추천 연구는 콘텐츠 필터링, 협업적 필터링, 지식기반 추천 기법을 활용하였으며, 구직자와 구인자를 동시에 만족시키는 알고리즘이 연구되었다 [3,4]. 나아가, 구직자의 과거 이직 기록을 사용하여 기계학습 기법으로 다음 직업을 예측하는 연구도 있었다 [5].

하지만 지금까지의 직업 추천 연구는 특정 연령을 대상으로 진행되지 않았다. 기존의 알고리즘[3,4]에서도 나이를 사용자 특성으로 반영하였으나, 중장년층의 세대 특성을 반영한 어느 특정 연령층 대상의 직업 추천 알고리즘은 미흡하다. 과거 중장년층을 대상으로 한 사용자 조사 연구에서 [6], 중장년층은 일자리를 찾을 때 과거의 직업에 따라 경력이나 자존심을 중요시하거나, 거주지와 직장이 적정 거리에 있으며 업무의 스트레스가 적은 직종을 선호하는 등 청년층의 구직 특성과는 다른

직종코드	2-depth	3-depth
4	영업,판매,운송관리자	
017100		영업 및 판매 관련 관리자
017101		영업 관련 관리자(영업소장 및 지점장)
017102		판매 관련 관리자(판매점장)
017200		운송 관련 관리자

표 1 워크넷 직종 분류표 예시

1) <http://www.work.go.kr/>

최상위 선호직종 순위	최상위 out-domain 순위		
	1순위	2순위	3순위
전기·전자 부품 및 제품 생산직	간판제작, 기타 생산직	제조, 건설, 농림어업 단순노무직	기계조립·금속 생산직
재무, 회계, 경리 사무직	비서, 사무보조, 전산자료입력, 설문조사	경영, 홍보, 인사, 총무 사무직	안내, 접수, 고객상담
간판제작, 기타 생산직	전기·전자 부품 및 제품 생산직	기계조립·금속 생산직	식품 가공 및 생산직
제조, 건설, 농림어업 단순노무직	전기·전자 부품 및 제품 생산직	기계조립·금속 생산직	식품 가공 및 생산직
경영, 홍보, 인사, 총무 사무직	재무, 회계, 경리 사무직	비서, 사무보조, 전산자료입력, 설문조사	무역, 운송, 자재, 생산·품질관리 사무직

표 2 20-40대 최상위 선호직종 (5개) 및 최상위 Out-domain 직종 (3개)

최상위 선호직종 순위	최상위 out-domain 순위		
	1순위	2순위	3순위
청소, 방역, 가사서비스, 기타서비스	경비·경호·보안	제조, 건설, 농림어업 단순노무직	판매, 서비스 단순직
경비·경호·보안	청소, 방역, 가사서비스, 기타서비스	판매, 서비스 단순직	계기검침, 수급, 주차, 기타 서비스
간판제작, 기타 생산직	청소, 방역, 가사서비스, 기타서비스	식품 가공 및 생산직	전기·전자 부품 및 제품 생산직
보건의료 단순종사자 (간병, 요양보호, 산후조리)	청소, 방역, 가사서비스, 기타서비스	조리사(주방장) 및 바텐더	웨이터, 홀씨빙, 주방보조, 요리배달
제조, 건설, 농림어업 단순노무직	청소, 방역, 가사서비스, 기타서비스	식품 가공 및 생산직	전기·전자 부품 및 제품 생산직

표 3 50-70대 최상위 선호직종 (5개) 및 최상위 Out-domain 직종 (3개)

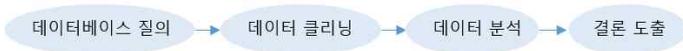


그림 1 데이터 분석 과정

것으로 파악되었으며, 이러한 특성을 직업추천에 반영할 필요성이 있다. 또한, 구인 회사나 구직자 대상의 대규모 설문이나 조사 방식의 연구는 있었지만, 이러한 접근 방식은 과도한 조사비용을 초래하며, 응답자의 상태가 과도하게 반영되거나, 연구자의 객관성이 떨어지는 단점이 있다. 따라서 실제 운영되고 있는 온라인 구직 서비스의 실제 데이터에 기반하여 중장년층의 직업 알선 트렌드를 파악하는 연구가 요구된다.

이에 따라 본 연구에서는 중장년층에 특화된 알고리즘을 개발하기 위하여, 워크넷 사이트의 실제 구직자 데이터(3년, 약 18만개의 기록)를 기반으로 중장년층 직업 추천 알고리즘에 반영할 수 있는 중장년층 특성분석을 목표로 한다. 구체적으로 연령대별로 사용자가 선호한 직종은 무엇이었는지 실제로 취업한 직종은 무엇이었는지 알아보고, 어느 정도의 비율로 취업하였는지에 관한 통계를 연령대 별로 분석한다. 직종은 워크넷 직종분류표를 기준으로 하였다. 세대별로 다른 특성을 파악한다면 연령대 별로 직업 추천 등에 이를 활용 할 수 있다.

## 2. 실험 데이터 설명

본 연구는 워크넷 웹사이트에서 2013년~2015년까지 3년에 걸쳐 수집된 182,516개의 데이터를 바탕으로 한다. 데이터는 MySQL 데이터베이스에 9개의 테이블로 존재하며 사용자 정보, 회사정보, 사용자 클릭정보 등을 포함한다. 따라서 본 연구에서 사용할 데이터를 추출하기 위해 질의 과정이 필요하며 질의를 통해 최종적으로 분석하게 될 데이터는 사용자 식별번호, 생년월일, 구직자가 입력한 1~3순위의 선호직종, 실제 취업한 직업의 선호직종 칼럼으로 구성된다.

또한, 본 연구에서는 워크넷에서 사용하는 취업 알선 직업 분류표를 사용한다. 이 분류표는 총 3개의 depth로 구성되며, 1-depth는 15개, 2-depth는 142개, 3-depth는 1904개의 직종 카테고리가 있으며 워크넷에서 사용되는 직업들을 이 기준에 의해 분류한다. 또한, 직종 분류표는 1-depth 안에 2-depth가 포함되고 2-depth 안에 3-depth가 포함되는 계층적 구조로 되어 있다.

편의상 2-depth의 직종코드에 0부터 130번까지의 임의 코드를 부여하였다. 표 1은 직종의 예를 보여준다. 1-depth는 본 연

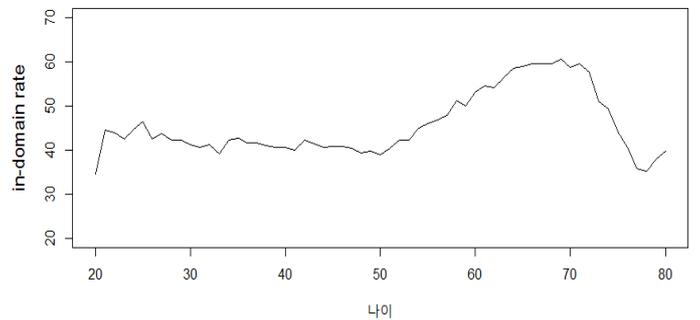


그림 2 나이대별(20-80세) In-domain 비율

구에서 사용하지 않아 표 1에 포함하지 않았다.

## 3. 분석 과정 및 방법

그림 1은 본 연구의 데이터 분석 과정을 도식화하여 보여준다. 데이터 분석에 앞서 데이터베이스에서 필요한 데이터를 추출하는 질의 과정이 필요하다. 또한 18만여 명의 대량 데이터를 다루기 위해서는 데이터 클리닝 과정이 필요하다. 구직자 데이터 중 선호 직종이 실제 직종 분류표에 존재하지 않는 것들이 존재했기 때문에 이러한 유저들을 제거해 줄 필요가 있었다. 또한 기준이 되는 직종 분류표에 중복되는 직종이 2-depth에 나타나는데 이러한 문제점을 해결해 주기 위해 직종 중복을 제거해주는 단계가 필요했다. 데이터 클리닝 결과 총 182516명의 구직자 중 선호직종 오류로 인해 10명의 구직자가 데이터에서 제거되었고, 직업 분류표에서 2-depth 직종의 개수가 142개에서 130개로 축소되었다.

본 연구에서는 구직자가 선호한 직종 카테고리과 실제 취업한 직종의 카테고리에 따라 In-domain, Out-domain 이라는 용어를 사용한다. 각각의 정의는 다음과 같다.

- **In-domain 취업:** 구직자가 선호한 직종(3-depth)과 최종적으로 구직자가 취업한 직종(3-depth)이 같은 2-depth 직종 카테고리에 포함되었을 시 In-domain 취업이라 한다.

- **Out-domain 취업:** 구직자가 선호한 직종(3-depth)과 최종적으로 구직자가 취업한 직종(3-depth)이 같은 2-depth 직종 카테고리에 포함되지 않았을 시 Out-domain 취업이라 한다.

위의 정의를 사용하여 연령대별 (20-80세)로 구직자의 In-domain 비율을 계산한다. 또한, 연령대 별로 가장 많이 선호

	1순위	2순위	3순위
20~40대	36	35	37
50~70대	18	27	35

표 4 최상위 Out-domain 직종 수

된 순으로 선호직종을 추출하고 각각의 직종에 따라 최상위 Out-domain 직종 3개를 추출한다. 예를 들어, 한 사용자가 79 (2-depth) 직종을 선호하였지만 80 (2-depth) 직종으로 최종 취업을 한 경우 80 (2-depth) 직종은 Out-domain으로 계산된다. 이와 같은 방법으로 79 (2-depth) 직종을 선택했었던 모든 사용자에 대해 Out-domain의 빈도수를 계산하여 가장 많은 Out-domain 빈도수를 가진 3개의 직종을 추출한다.

4. 분석 결과

4.1 나이 별 In-domain 분석 결과

그림 2는 x축이 나이, y축이 In-domain(백분율) 비율을 나타내며 나이 별로 In-domain 비율을 그래프로 표현한 것이다. 그래프에서 볼 수 있듯이 55-75세 사이의 구직자 In-domain 비율(54.16%)이 20-54세 구직자 In-domain 비율(41.56%)에 비해 약 13%가량 높았다. 중장년층 구직자가 본래 원하던 직종으로 혹은 본래 원하던 직종과 비슷한 직종으로 갈 확률이 젊은 층보다 높다는(30대: 41.25%, 60대: 57.51%) 결과를 볼 수 있다. 이러한 결과가 나타나게 된 원인을 다음 3가지로 추측할 수 있다. 첫째, 중장년층 구직자는 자신의 한계 혹은 능력을 젊은 층보다 신중히 고려하여 선호직종을 선택했을 가능성이 있다. 둘째, 중장년층 구직자가 선택할 수 있는 직종이 제한적이어서 제한적인 특정 직종을 선택할 확률이 높고 그 직종으로 취업할 확률도 상대적으로 높았을 가능성이 있다. 셋째, 젊은층의 In-domain 비율이 상대적으로 낮는데, 이는 젊은층의 구직자는 선호직종이 다르더라도 다른 직업의 조건이 맞으면 그 직업을 선택할 가능성이 장년층보다 상대적으로 높았을 가능성이 있다.

4.2 연령대별 선호직종 및 Out-domain 직종

표 2와 3은 각각 청년층 (20~40대), 중장년층(50~70대)에 대한 연령대별 최상위 선호직종 (5개)과 각 직종의 최상위 Out-domain 취업직종들 (3개)의 예를 보여준다. 각 표의 첫 번째 열은 각 연령대에서 가장 많이 선호된 상위 직종 (2-depth)을 나열한 것이다. 2~4번째 열은 각 직종별로 상위 3개의 Out-domain 직종 (2-depth)을 나열한 것이다. 이 결과를 보면 20~40대 구직자들이 가장 많이 선호하는 직종은 생산직과 사무직이었으며 대안적으로 선택하는 직종도 다양하다는 것을 알 수 있다. 반면 50~70대 구직자들이 가장 많이 선호하는 직종은 단순 노무직 관련 직종들이었으며 (청소, 방역, 가사서비스; 경비 보안 등) 실제로 취업한 직종도 단순 노무직 관련 직종들 (경비, 보안; 청소, 방역, 가사서비스 등)이었음을 알 수 있다. 특히 50~70대의 Out-domain 순위 1의 결과를 보면, 다양한 선호 직종이 있더라도 최종 취업 직종이 청소, 방역, 가사서비스에 집중되어 있음을 알 수 있다. 이를 미루어 보아 대한민국의 중장년층들은 단순 노무직을 가장 많이 선호하고, 선택한 직종에 취업하지 못했을 시 청소 관련 단순 노무 직종에 대안적으로 취업 하고 있음을 알 수 있다.

표 4는 청년층과 중장년층의 Out-domain 직종 카테고리를 중복을 제거한 후 직종의 개수를 나타낸 것이다. 이 표에서 확인할 수 있듯이 나이가 상대적으로 젊은 층의 구직자들은 자신이 선호한 직종 이외에 다른 직종으로 취업을 할 경우 중장년층에 비해 다양한 직종을 선택하여 취업함을 알 수 있다. 즉,

나이가 많은 중장년층일수록 자신이 원하는 직종이 아닌 다른 직종으로 취업을 할 경우 취업할 수 있는 직종이 제한적이거나 선택의 폭이 넓지 않음을 보여준다.

4.3 직업 추천시스템 디자인 함의 (Design Implication)

본 연구의 분석 결과에서 볼 수 있듯이 In-domain 비율이 중장년층에서 더 높게 나타난 것은 중장년층 구직자가 젊은 층 구직자에 비해 선호 직종에 더 높은 우선순위를 준다는 것으로 해석할 수 있을 것이다. 이 결과는 향후 직업 추천 알고리즘 개발 시에 나이를 선호직종에 대한 가중치 판별에 사용할 수 있을 것이다. 또는 구직자가 선호 직종과 더불어 다른 선호 요 구사항(임금, 거리 등)들을 가지고 있을 때에 젊은 층의 구직자에게는 다른 조건들을 바꿔서 추천결과를 보여주는 것 보다 선호직종을 바꾸는 것이 더 만족스러운 결과를 줄 수 있고, 중장년층의 경우에는 선호직종보다 다른 조건들을 변화시키는 것이 더 만족스러운 결과를 줄 수 있을 것이라고 예측할 수 있다.

또한 두 번째 분석결과는 향후 직업 추천 알고리즘 개발 시 나이를 다양성(diversity)을 결정하는 요소로 사용할 수 있는 근거를 제시할 수 있다. 결과에서 보았듯이 젊은 층은 자신이 원하는 직종이 아닌 다른 직종으로 취업할 시에 중장년층보다 다양한 직종을 선택하는 것을 볼 수 있었다. 즉 중장년층이 젊은 층에 비해 다양하지 않은 직종들로 취업함을 볼 수 있었다. 이것은 추천 알고리즘에서 고려되어야 할 다양성 문제를 해결하는데 기여할 수 있을 것이다.

5. 결론 및 향후 연구

본 연구는 중장년층 구직자를 위한 직업추천 알고리즘의 개발을 목적으로, 정부에서 운영하는 워크넷에서 보유한 18만 건 이상의 실제 데이터에 기반하여 수행되었다.

본 연구에서는 최초 선호직종과 최종 취업 직종 사이의 관계만을 고려하여 수행되었지만, 앞으로의 연구에서는 다양한 특징들을 고려하여 중장년층 구직자들만이 가진 특성을 분석하게 될 것이다. 특히 다양성 판별과 관련하여서는 워크넷에서 3순위 위까지 선호 직종을 입력하는 특성을 고려하여 1순위 선호 직종 뿐 아니라 2, 3순위 선호 직종을 함께 분석에 사용함으로써 더욱 고도화된 다양성 판별을 할 수 있을 것이다. 또한 직업의 In-domain 비율을 나이에만 국한되어 분석하지 않고, 성별, 학력 등의 관점으로 분석한다면 더욱 고도화된 분석 결과를 얻을 수 있을 것이다.

참 고 문 헌

[1] 통계청, “2015 고령자 통계”, 2015.  
 [2] 워크넷, <http://www.work.go.kr/empSpt/statsData/labStats.do>, 2015.  
 [3] Siting Z., Wenxing H., “Job recommender systems: A survey.”, 7th International Conference on Computer Sci, 2012.  
 [4] Hong W., Li L., Li T., & Pan W.,- “iHR: an online recruiting system for Xiamen Talent Service Center.-”, Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 1177-1185,-2013.  
 [5] Paparrizos I., Cambazoglu B. B., & Gionis A.,- “Machine learned job recommendation.-” Proceedings of the Fifth ACM Conference on Recommender Systems - RecSys ' 11, 325, 2011.-  
 [6] 장진철, 조대봉, 김성찬, 진효진, 김성중, 이문용, “중장년층을 위한 개인화 일자리 매칭서비스 요구사항 분석”, 한국HCI학회, 2016.