

노동정책연구
2012. 제12권 제3호 pp.25-51
한국노동연구원

연구논문

구인·구직조건을 유연성을 고려한 잡매칭함수의 개발

박성익*
조장식**
전현중***
김종한****
류장수*****

현재 워크넷의 소프트매칭(soft matching)은 희망 근로조건 입력 후 검색되어 나타난 직업의 구직자 적합도를 계산하는 데 필요한 연관성 값을 자의적으로 설정하고 있다. 본 논문은 합리적이고 현실적인 연관성 값을 계산하기 위해서, 워크넷(구직표, 구인표)과 고용보험 DB를 연계하여 주민 특성과 업체 특성 간의 일치도 확률분포를 파악함으로써 잡매칭함수를 개발하였다. 잡매칭함수는 1단계에서 구직자에게 적합한 중분류 직종을 선정하고, 2단계에서 적합한 기업을 선정하는 단계별 접근방법을 취하였다. 주요 결과는 다음과 같다. 첫째, 1단계 직종선정을 위한 직종별 로지스틱 분석결과와 설명력이 양호하여, 구직자별로 적절한 직종을 선택하는 것이 가능하다는 결과를 도출하였다. 둘째, 2단계 세분류 직종 및 기업 선정 연구결과에 의하면, 특성차이별로 도출된 확률분포가 일반적인 예상과 부합될 뿐만 아니라 산출된 잡매칭함수의 현실 설명력이 높은 것으로 평가된다.

핵심 용어 : 소프트매칭, 잡매칭함수, 워크넷, 고용보험 DB

논문접수일: 2012년 5월 8일, 심사의뢰일: 2012년 6월 15일, 심사완료일: 2012년 8월 10일

* (제1저자) 경성대학교 국제무역통상학과 교수(sipark@ks.ac.kr)

** 경성대학교 통계학과 교수(jscho@ks.ac.kr)

*** 동서대학교 국제통상학과 교수(juss90@hanafos.com)

**** 경성대학교 경제물류학부 교수(jhkim@ks.ac.kr)

***** (교신저자) 부경대학교 경제학부 교수(jsryu@pknu.ac.kr)

I. 서론

현재 워크넷을 통한 취업알선의 성공률은 극히 낮은 수준이다.¹⁾ 이처럼 취업알선의 성공률이 낮은 것은, 적합한 구직자나 기업을 우선 선별하지 않고 정보의 홍수 속에서 너무 많은 정보가 유통됨에 따라, 주민은 주민대로 기업을 선별하기 어렵고 기업은 기업대로 구직자 정보를 스팸메일 정도로 치부하는 상황에 기인한다. 따라서 성공적인 잡매칭을 위해서는 무조건 많은 양의 정보가 필요한 것이 아니라 정보탐색비용을 최소화하면서 기업과 주민에게 적합한 맞춤형 정보를 제공하는 것이 필요하다.

이러한 인식하에 현재 워크넷은 노동수급 당사자 간의 희망 고용조건이 완전히 일치될 때만 검색결과를 보여주는 방법(hard matching)과 그와 병행하여 일부 조건만 충족하더라도 연관된 검색결과를 보여주는 매칭(soft matching)을 동시에 활용하고 있다. 소프트매칭은 직업 간 유사성이나 상관관계 및 입력된 검색 조건의 중요도에 따라 사용자의 관심이나 희망사항을 유연하게 반영하여 잡매칭 가능성을 높이는 검색결과를 제공한다는 측면에서 하드매칭보다 바람직하다. 현실적으로 고용수급자 간의 고용조건이 100% 일치하지 않더라도 잡매칭은 가능할 수 있으므로 미스매칭의 문제를 해결하여 취업률을 제고하는 데는 소프트매칭이 더 바람직한 것이다.

그러나 현재 워크넷에서 적용하고 있는 소프트매칭의 단점은 희망 근로조건 입력 후 검색되어 나타난 직업의 구직자 적합도를 계산하는 데 필요한 연관성 값이 자의적으로 주어진다는 점이다. 예를 들어, 현재 워크넷 소프트매칭은 구직자가 240만 원을 요구하고 기업이 230만 원을 제시할 때 연관성을 50%로 설정하고 있다. 아울러 구직자와 직업의 연관성을 전체적으로 계산할 때 거리나 직종 그리고 임금을 전부 동일한 비중으로 계산하고 있다.²⁾ 이와 같이 현재

1) 본 연구에서 산출한 2009년 7월 1일부터 2010년 6월 30일까지 직업알선의 성공률은 1.4%에 불과하다.

2) 김형래(2010), 고용촉진지구 주관 워크숍.

의 연관성 값은 합리적으로 설정되어 있다고 보기 어렵다.

합리적이고 현실적인 연관성 값을 계산하기 위해서는 워크넷과 고용보험에 정보가 수록되어 있는 구직자가 어떤 직업에 취업하였는지를 분석할 필요가 있다. 워크넷과 고용보험 DB를 연계하여 취업자의 특성을 분석하면, 구직자나 구인자가 희망하는 조건이 전부 충족되어야만 취업이 이루어지는 것이 아니라는 점을 확인할 수 있다. 이로부터 구직자와 구인자가 감내하는 수준의 근로조건 및 희망채용조건에 대한 확률분포를 파악할 수가 있다. 이러한 확률분포를 활용하게 되면 현실적인 연관성 지수, 즉 근로조건과 채용조건에 대한 가중치를 파악할 수가 있게 되는 것이다.

이러한 인식하에 본 논문은 워크넷(구직표, 구인표)과 고용보험 DB를 연계하여 주민 특성과 기업 특성 간의 일치도 확률분포를 파악함으로써 잡매칭함수를 개발하는 것을 목표로 한다.

본 논문의 순서는 다음과 같다. 우선 제Ⅱ장에서는 근로조건 유연성과 소프트웨어매칭에 관련된 선행연구를 간략히 검토한다. 그리고 제Ⅲ장에서는 워크넷(구직표)과 고용보험 DB를 연계하여 직종별 구직자의 특성을 살펴보고 구직자가 희망하는 직종에 취업하였는지를 분석한다. 제Ⅳ장에서는 한국고용직업분류표의 중분류 직종별로 구직자에게 적합한 직종을 선정하는 과정을 설명한다. 덧붙여 직종코드 155(금형 및 공작기계 조작용)를 대상으로 직종 및 기업 선택 과정을 단계별로 분석한다. 그리고 마지막 제Ⅴ장에서 본 연구의 결과를 정리하고 한계를 언급한 후 향후의 연구방향을 제시한다.

Ⅱ. 선행연구

잡매칭이론(job-matching theory)은 크게 보아 개별 구직자의 미시경제학적 의사결정을 연구하는 탐색이론(search theory)과 다양한 형태의 구직자 간 상호작용이 가져오는 거시경제학적 결과를 연구하는 이론(matching theory)으로 구분할 수 있다. 거시경제학적 측면에서는 Mortensen and Pissarides(1994)가 처음으로 매칭이론의 개념을 도입한 이래, Pissarides(2000), Petrongolo and

Pissarides(2001), Shimer(2005), Eyiungor(2010) 등이 임금과 실업 그리고 경기변동이 어떻게 결정되는지에 대하여 연구하였다. 이에 비하여 우리나라에서는 취약계층의 고용에 관련하여 미시적인 연구도 많이 이루어졌다. 특히 취약계층의 고용에 관해서는 주로 취약계층의 실태 분석, 직업능력개발방안, 금전적 인센티브(예: 고용보조금)에 대한 효과, 고용과 복지를 연계하는 방안 등에 대해서 연구가 진행되었다. 그러나 실질적으로 이들 취약계층이 일자리 탐색에 있어서 얼마나 유연한 입장을 가지고 있으며, 취약계층이 일자리를 가질 수 있는 지역의 중소기업들이 얼마나 취약계층의 구인에 대해 유연한 입장을 가지는지에 대한 연구는 미흡한 실정이다.

취약계층의 구직 유연성과 중소기업의 구인 유연성에 대한 연구는 부분적으로 진행되었다. 대표적으로 이장원(2001)의 연구에서는 근로취약계층의 경우 “힘든 일이라도 하겠다”는 강한 근로 동기를 가진 계층은 구직활동과 직업훈련에 더 적극적임을 지적하였다. 김성안 외(2005)는 충청지역의 취약계층을 대상으로 고용증진 방안을 연구하였다. 고령자, 여성가장, 장애인을 대상으로 한 사례조사 분석결과에 따르면, 근로조건 유연성은 고령자가 가장 유연하며, 다음이 장애자, 마지막으로 여성이었다.

다음으로 중소기업의 구인 유연성에 관해서는 송호근·설동훈(2001)의 연구가 있다. 이 연구에서는 유연성을 크게 연령, 성별, 학력, 경력으로 나누어서 분석하였다. 이들은 연령에 가장 유연한 직종은 생산기능인력(76.7%)이었으며, 성별에 대해서는 사무관리인력(67.7%)이 가장 유연하다는 결과를 발견하였다. 그리고 학력에 대한 유연성은 생산기능인력(87.7%)이 다른 직종에 비해 높았다. 마지막으로 경력에 관해서는 사무관리인력이 64.9%로 가장 유연성이 높다는 결과를 제시하였다.

한편 조인호(2009)는 워크넷의 구인·구직시스템이 하드매칭방법을 적용하여 검색조건이 복잡해지면 해당 구인정보가 안 나오는 문제가 존재한다는 지적을 하고 있다. 아울러 새로운 직업이 나타나거나 정확한 직업 명칭을 모르는 경우에는 해당 구인정보를 제공하지 못할 수도 있다는 문제점을 지적하고 있다. 그리고 오성욱(2009)은 스웨덴의 인터넷 매칭 취업서비스를 소개하고 있다. 그의 연구에서는 고용서비스 기능 강화를 위한 신기술(소프트매칭) 도입 및 운영 부

분을 설명하고 있다.

최근 들어 이와 같이 일자리 매칭에 있어서 하드매칭보다는 소프트매칭을 활용해야 할 필요성이 강조되고 있다. 소프트매칭은 사용자의 심리가 반영된 중요도가 높은 조건과 항목별 상관관계가 높은 정보에 가중치를 부여함으로써, 사용자의 의도에 가장 근접한 결과를 찾아주는 방법이다. 소프트매칭의 장점으로서는 항목에 대한 중요성과 조건(유사직종, 학력, 거리 등)에 대한 허용(tolerance) 정도에 관한 상담원의 경험과 노하우(Know-how)를 시스템에 반영할 수 있다는 점이다. 즉 소프트매칭에서는 구직자가 반경 2km 내의 일자리를 찾는다고 하여도 임금이 높다면 4km 거리의 일자리도 구직자에 따라 수용이 가능함을 의미한다.

이와 같이 많은 연구가 소프트매칭의 중요성을 강조하고 있지만, 중요도 높은 조건에 대하여 어떻게 가중치를 부여해야 하는지에 대해서는 아직 연구가 이루어진 것이 없다. 예외적으로 박성익·전현중 외(2009)가 잡매칭함수를 개발하여 그 적용 가능성을 연구한 바가 있다. 그러나 그들의 잡매칭함수는 성, 연령과 같은 모든 특성이 동일한 정도로 중요한 것이 아님에도 불구하고 전부 동일한 가중치를 부여하고 있다는 문제점이 있다. 그리고 임계치의 값이 자의적이다. 예를 들어 구직자나 구인자가 어느 정도의 임금 차이를 수용할 수 있는지를 알 수가 없다. 또한 그들은 구직자가 해당 직종에 근무하기를 원한다는 것을 선결조건으로 하여 해당 직종에서의 잡매칭함수 값을 계산하였으나, 특정 직종을 희망한다고 해서 그 직종에 반드시 채용되는 것은 아니라는 점을 간과하였다.

이러한 문제점을 주목하여, 본 논문은 워크넷과 고용보험 DB를 연계하여 구직자나 구인자가 희망근로조건이나 희망채용조건에서 얼마나 유연한지를 통계적으로 분석하도록 한다. 또한 본 연구에서는 1단계의 직종선택, 2단계의 기업선택의 두 단계로 나누어 적합한 직장을 선정하도록 한다. 희망직종과 실제직종이 반드시 일치하는 것은 아니지만, 그럼에도 불구하고 구인자는 직종 단위로 구인을 하는 것이 관행이므로 구직자에게 적합한 직종을 먼저 선정하는 것이 바람직하다고 할 수 있다. 또한 직종을 먼저 선정하지 않고 모든 주민의 모든 기업에 대한 매칭함수 값을 계산하여 적절한 직업을 알선하고자 할 경우,

그에 따른 계산상의 부담이 상당할 것으로 판단된다. 따라서 본 논문에서는 잡매칭과정을 2단계로 구분³⁾하여 구직자에게 직종 및 기업을 추천하는 과정을 거치도록 한다.

Ⅲ. 워크넷(구직포)과 고용보험 DB를 연계한 분석

1. 직종별 취업자 특성

본 논문은 2009년 7월 1일부터 2010년 6월 30일까지 1년 동안 워크넷에 상용직 구직을 등록한 부산지역 거주자 94,359명을 대상으로 분석하였다. 이들의 취업 현황을 살펴보면, 전체 94,359명 중에서 42,467명(45.0%)이 취업을 하고 51,892명(55.0%)이 취업을 하지 못했던 것으로 나타났다. 그런데 42,467명 중에서 42,331명에 대해서만 취업한 직종이 확인되었으므로, 이들만을 대상으로 분석하였다.

인적특성별·직종별 취업자 특성은 <표 1>에 정리되어 있다. 우선 남자의 비중이 높은 직종은 9(96.3%), 15(96.1%), 16(93.9%)의 순이다. 남자의 비중이 낮은 직종은 7(5.6%), 6(12.1%), 13(19.1%)의 순이다. 성별에 따른 직종별 취업자 비중에 비교적 큰 차이가 있는 것으로 보인다.

청년층의 비중이 높은 직종은 3(66.9%), 8(65.0%)이고 낮은 직종은 23(4.3%), 11(5.2%)이다. 30대의 비중이 높은 직종은 15(39.8%), 20(39.7%), 16(39.0%)이고 낮은 직종은 11(5.9%)이다. 40대의 비중이 높은 직종은 18(34.7%), 13(34.2%)이고 낮은 직종은 8(3.1%)와 20(4.5%)이다. 50대의 비중이 높은 직종은 11(42.3%), 23(36.1%)이고 낮은 직종은 5(1.3%)와 3(2.0%), 20(2.0%)이다. 그리고 60세 이상의 비중이 높은 직종은 11(36.6%), 23(25.8%)이고 낮은 직종은 3(0.0%),

3) 잡매칭과정을 2단계로 구분하는 또 다른 이유는 직종 선택에서 중요한 역할을 하는 성 및 연령 정보를 법적 규제로 인하여 기업선정단계에서 활용할 수 없기 때문이다. 따라서 1단계에서는 성 및 연령 정보를 같이 활용하여 직종을 선정하고, 그 다음 단계에서 기업을 선정하는 단계로 직업 알선 단계를 구상하였다.

20(0.4%), 8(0.7%)과 10(0.7%)이다. 연령에 따른 직종별 취업자 비중에도 비교적 큰 차이가 있는 것으로 판단된다.

〈표 1〉 직종별 인적특성

(단위: 명, %)

	취업자	성별		연령					학력			
		남	여	15~29	30~39	40~49	50~59	60 이상	중졸 이하	고졸 이하	전문대졸 이하	대졸 이상
전체	42,331	54.7	45.3	34.7	30.7	15.7	12.8	6.1	11.5	32.4	26.7	29.4
1	5,319	61.8	38.2	39.8	36.8	13.5	7.4	2.6	4.1	20.1	32.1	43.7
2	10,045	30.0	70.0	51.4	33.6	10.0	3.8	1.2	2.2	23.1	33.1	41.6
3	405	22.8	77.2	66.9	25.2	5.9	2.0	0.0	0.5	11.6	25.7	62.2
4	491	28.5	71.5	43.0	37.1	14.3	4.5	1.2	0.0	7.3	20.2	72.5
5	80	46.3	53.8	51.3	38.8	7.5	1.3	1.3	0.0	12.5	31.3	56.3
6	1,903	12.1	87.9	40.8	28.1	17.7	11.1	2.3	7.3	33.8	38.0	21.0
7	1,326	5.6	94.4	24.3	30.8	27.6	14.8	2.5	6.6	29.0	35.7	28.7
8	286	36.7	63.3	65.0	30.1	3.1	1.0	0.7	1.4	9.1	32.9	56.6
9	1,627	96.3	3.7	20.7	37.6	19.9	16.5	5.4	9.2	47.9	23.8	19.0
10	1,692	55.8	44.2	39.5	35.7	20.0	4.1	0.7	2.4	32.9	29.9	34.8
11	3,059	61.0	39.0	5.2	5.9	9.9	42.3	36.6	46.5	38.7	6.8	8.0
12	280	34.3	65.7	56.8	21.8	10.7	8.6	2.1	6.1	19.6	29.3	45.0
13	1,000	19.1	80.9	19.2	18.4	34.2	25.9	2.3	18.3	49.9	17.5	14.3
14	1,476	74.9	25.1	16.3	26.5	17.3	24.3	15.7	28.1	24.3	18.9	28.7
15	1,876	96.1	3.9	33.0	39.8	16.2	9.2	1.9	6.2	41.5	30.1	22.1
16	1,002	93.9	6.1	30.8	39.0	17.7	10.9	1.6	9.5	51.2	23.9	15.5
17	161	82.6	17.4	32.3	37.3	15.5	13.7	1.2	9.3	37.3	23.0	30.4
18	245	47.5	52.5	18.8	18.0	34.7	24.5	4.1	31.8	44.9	16.3	6.9
19	1,390	83.9	16.1	39.4	37.1	15.5	6.6	1.4	3.5	42.6	31.5	22.4
20	715	84.3	15.7	53.4	39.7	4.5	2.0	0.4	0.0	15.4	34.4	50.2
21	140	49.3	50.7	32.1	31.4	23.6	10.7	2.1	11.4	36.4	27.9	24.3
22	7,296	71.0	29.0	24.9	29.3	21.0	17.3	7.5	18.8	45.3	20.7	15.3
23	512	74.8	25.2	4.3	11.3	22.5	36.1	25.8	44.9	42.2	2.3	10.5

주: 직종분류: 01 관리직, 02 경영회계사무관련직, 03 금융보험관련직, 04 교육 및 자연과학사회과학연구관련직, 05 법률경찰소방교도관련직, 06 보건의료관련직, 07 사회복지 및 종교관련직, 08 문화예술디자인방송관련직, 09 운전 및 운송관련직, 10 영업 및 판매관련직, 11 경비 및 청소관련직, 12 미용숙박여행오락스포츠관련직, 13 음식서비스관련직, 14 건설관련직, 15 기계관련직, 16 재료관련직, 17 화학관련직, 18 섬유 및 의복관련직, 19 전기·전자관련직, 20 정보통신관련직, 21 식품가공관련직, 22 환경인쇄목재가구공예 및 생산관련직, 23 농림어업관련직

중졸 이하의 비중이 높은 직종은 16(46.5%), 23(44.9%)이고 낮은 직종은 4(0.0%), 5(0.0%), 20(0.0%)이다. 고졸 이하의 비중이 높은 직종은 16(51.2%), 13(49.9%), 9(47.9%)이고 낮은 직종은 4(7.3%), 8(9.1%)이다. 전문대졸 이하의 비중이 높은 직종은 6(38.0%)이고 낮은 직종은 23(2.3%)과 11(6.8%)이다. 특히 전문대졸 이하의 경우는 평균비중 26.7%에 비하여 표준편차 9.0%가 작은 편이어서, 다른 학력에 비하여 특정 직종에 크게 편중되지 않고 고르게 취업하였음을 알 수 있다. 그리고 대졸 이상의 비중이 높은 직종은 4(72.5%), 3(62.2%)이고 낮은 직종은 18(6.9%), 11(8.0%)이다. 학력에 따라서도 직종별 취업자 비중에 비교적 큰 차이가 있는 것으로 생각된다.

2. 희망 및 실제 직종별 취업 여부

워크넷에는 본인이 희망하는 직종을 1, 2, 3 순위로 세 개의 직종까지 5자리 단위로 기록할 수 있게 되어 있다. 그러나 고용보험 DB의 정보에서는 취업 직종을 3자리까지만 확인할 수 있다. 따라서 희망직종 1, 2, 3순위 중 어느 하나만이라도 실제직종과 3자리까지만 일치하면 직종이 일치하는 것으로 판정하도록 하였다. 그리고 희망하는 직종에 ‘상관없음’이라고 답변한 취업자의 경우는 희망직종이 실제직종과 일치한 것으로 판정하였다.

분석결과, 희망 및 실제 직종이 일치한 취업자는 전체 42,331명 중에서 23,652명(55.7%)인 것으로 나타났다. 그 중 1순위가 충족된 사람은 19,819명(46.8%), 2순위가 충족된 사람은 2,962명(7.0%), 3순위가 충족된 사람은 871명(2.1%)인 것으로 밝혀졌다. 희망직종이 실제직종과 일치하는 취업자의 비중이 55.7%로 높게 나타났다. 그 이유는 구직자의 특정 직종 희망여부가 해당 직종에 대한 구직자의 경험이나 능력 및 의사 등을 반영하는 유용한 대리변수(proxy variable)로서 해당 직종 취업에 긍정적인 역할을 하기 때문인 것으로 생각된다.

그러나 이와 동시에 희망직종과 실제직종이 일치해야만 취업을 하는 것은 아니라는 점을 분명히 보여주고 있다. 즉 직업알선에 있어서 소프트웨어 매칭이 중요하다는 점을 다시 한 번 확인할 수 있다. 그러므로 구직자에게 적합한 직종을 선택하는 데 있어서 구직자의 희망 여부만을 고려할 것이 아니라 다른 요인도

같이 고려하여 적합한 직종을 선정할 필요가 있다고 하겠다.

1순위만을 기준으로 희망 및 실제 직종별 취업 결과를 분석한 결과는 <표 2>에 정리되어 있다. 구직자 중에서는 가장 많은 25,573명(27.1%)이 직종 2에 취업을 희망했던 것으로 나타났다. 그 다음은 직종 22의 8,101명(8.6%), 직종 11의 7,214명(7.6%), 직종 15의 5,914명(6.3%) 순인 것으로 나타났다. 이에 반하여 직종 5의 213명(0.2%), 직종 17의 510명(0.5%), 직종 3의 660명(0.7%), 직종 21의 724명(0.8%) 순으로 적은 인원이 해당 직종에 취업을 희망했던 것으로 나타났다.

실제 취업자의 구성을 보면, 직종 2의 10,045명(28.6%)이 가장 많은 것으로 나타났으며, 그 다음은 직종 15의 1,876명(8.0%), 직종 22의 7,296명(7.9%)의 순인 것으로 나타났다. 이에 반하여 직종 5의 80명(0.2%), 직종 23의 512명(0.5%), 직종 17의 161명(0.6%), 직종 3의 405명(0.7%)의 순으로 적은 인원이 해당 직종에 취업한 것으로 나타났다.

<표 2>에 따르면 실제 취업자의 직종 분포는 희망했던 직종의 분포와는 다른 모습을 보여주고 있는 것을 알 수 있다. 대표적인 직종으로 직종 1(관리직)은 3,261명이 취업을 희망했으나 실제로는 5,319명이 취업을 하여 희망자 대비 실제 비율이 163.1%에 달하는 것으로 나타났다. 그리고 직종 22(환경인쇄목재가구공예및생산관련직)는 8,101명이 취업을 희망하였는데 실제로는 7,296명이 취업을 하여 그 비율이 90.1%에 달하는 것을 알 수 있다. 이에 비하여 직종 8(문화예술디자인방송관련직)은 3,522명이 취업을 희망하였으나 오직 286명만이 취업에 성공하여 희망자 대비 실제 비율이 8.1%로 대단히 낮은 것을 알 수 있다. 이와 같은 결과로 볼 때, 많은 구직자가 희망하는 직종에 취업을 하는 것이 아니라 희망하지 않는 직종에 취업을 하게 되는 것을 알 수 있다.

그런데 특정 직종에 취업한 사람이 반드시 그 직종에 취업을 희망했던 사람은 아닐 수 있다는 점을 유념할 필요가 있다. 이러한 점을 감안하여 희망자 중 실제 취업자를 나타낸 결과를 볼 필요가 있다. 우선 원하는 직종에 취업을 한 취업자 비율은 직종 7(사회복지및종교관련직)에서 37.5%로 가장 높고, 다음이 직종 6(보건의료관련직)이 31.8%, 직종 11(경비및청소관련직)이 31.0%의 순으로 높은 것을 알 수 있다. 이들 직종 중 7과 6에서 희망자의 취업확률이 높은

것은 특정 자격증이 필요하므로 그에 따른 준비가 되어 있는 구직자만 실제 취업에 성공하기 때문인 것으로 판단된다. 그리고 직종 11의 경우는 특별한 준비는 필요가 없지만, 3D업종으로 취업이 비교적 용이하기 때문인 것으로 판단된다. 이에 반하여 직종 8(문화예술디자인방송관련직)은 5.6%, 21(식품가공관련직)은 5.9%, 12(미용숙박여행오락스포츠관련직)는 6.0%로 취업확률이 매우 낮은 것으로 나타났다.

<표 2> 취업자의 희망 및 실제 직종별 취업 결과

(단위: 명, %)

	구직자		실제 취업자		실제취업자/ 희망자 비율 (=B/A*100)	1순위직종 희망자 중 실제 취업자		
	인원(A)	비중	인원(B)	비중		인원	희망자 중 비중	실제취 업자중 비중
전체	94,359	100.0	42,331	100.0	45.0	19,819	21.0	46.7
1	3,261	3.5	5,319	3.5	163.1	380	11.7	7.1
2	25,573	27.1	10,045	28.6	39.3	7,303	28.6	72.7
3	660	0.7	405	0.7	61.4	112	17.0	27.7
4	1,447	1.5	491	1.2	33.9	213	14.7	43.4
5	213	0.2	80	0.2	37.6	29	13.6	36.3
6	4,399	4.7	1,903	4.5	43.3	1,397	31.8	73.4
7	2,239	2.4	1,326	2.6	59.2	839	37.5	63.3
8	3,522	3.7	286	3.9	8.1	198	5.6	69.2
9	4,681	5.0	1,627	5.0	34.8	911	19.5	56.0
10	4,885	5.2	1,692	4.5	34.6	522	10.7	30.9
11	7,214	7.6	3,059	7.1	42.4	2,238	31.0	73.2
12	1,577	1.7	280	1.2	17.8	94	6.0	33.6
13	4,098	4.3	1,000	2.9	24.4	661	16.1	66.1
14	5,216	5.5	1,476	5.2	28.3	663	12.7	44.9
15	5,914	6.3	1,876	8.0	31.7	929	15.7	49.5
16	2,679	2.8	1,002	2.9	37.4	355	13.3	35.4
17	510	0.5	161	0.6	31.6	41	8.0	25.5
18	1,062	1.1	245	0.9	23.1	80	7.5	32.7
19	3,868	4.1	1,390	4.8	35.9	729	18.8	52.4
20	1,802	1.9	715	2.3	39.7	416	23.1	58.2
21	724	0.8	140	0.7	19.3	43	5.9	30.7
22	8,101	8.6	7,296	7.9	90.1	1,596	19.7	21.9
23	714	0.8	512	0.5	71.7	70	9.8	13.7

주: 직종분류는 <표 1>의 주를 참고

IV. 잡매칭함수의 설계 및 결과 분석

1. 잡매칭함수 설계방법

1단계에서는 구직자에게 적합한 직종을 선정하고, 2단계에서는 잡매칭(job matching)함수를 설계한 후 그 결과를 분석하도록 하였다. 그런데 직종 중분류의 경우는 모든 직종(1~23, 군인은 제외)에 대해서 취업확률을 추정하였다. 그러나 직종 세분류의 경우는 1개의 직종 세분류에 한해서만 취업확률을 추정하였다.

가. 직종선택

1단계의 직종 선택은 워크넷과 고용보험 DB의 객관적인 자료를 활용하여 분석하였다. 추정에는 모수추정방법(parametric method)인 로지스틱(logistic) 함수를 이용하여 추정하는 방법과 비모수추정법(non-parametric method)인 신경망(neural network) 분석방법을 이용하였다.

1단계에서 두 모형을 추정하는 데 활용되는 종속변수 및 독립변수는 전부 동일하다.⁴⁾ 단지 추정방법만 상이하다. 로지스틱 모형은 다음과 같다

$$y_i = \alpha_0 + \alpha_1 성_i + \alpha_2 연령_i + \alpha_3 학력_i + \alpha_4 자격증_i + \alpha_5 경력_i + \alpha_6 해당직종희망여부_i + \alpha_7 희망월급_i + \alpha_8 희망근무지_i + \epsilon_i \quad (1)$$

y_i = 해당직종 취업여부(0=미취업, 1=취업), i = 구직자 개인

ϵ_i = 오차항

신경망 모형은 신경세포의 작동방식을 모방하여 분석하는 것으로, 과거에 수집된 자료로부터 반복적인 학습과정을 거쳐 자료에 내재되어 있는 패턴을 찾아

4) 구직탐색이론에서 구직자가 직업 탐색 시 제안받는 임금 외에도 정규직 여부, 근무시간 유형 등은 중요한 고려사항이다. 그러나 자료의 한계로 인하여 포함시키지 못하였다.

내는 분석기법이다. 본 연구에서는 신경망 모형 중 가장 많이 사용되는 MLP (Multilayer Perceptron) 신경망⁵⁾을 활용하였다.

우선 두 모형에서 워크넷과 고용보험 DB의 자료를 이용하여 직종별로 사후적인 취업확률(\hat{y}_i)을 추정하였다. 그러면 특정 구직자의 \hat{y}_i 이 클수록 그 직종에 취업될 확률이 높다는 것을 의미한다. 그리고 추정된 모형의 예측력을 검증하고 가장 실제와 부합하는 모형을 선택한다. 이제 어느 구직자가 직업알선을 요청하면, 그 구직자의 특성 및 희망 값을 입력하여 \hat{y}_i 을 계산한다. 이 값이 큰 것으로 나타나는 직종이 구직자에게 적합한 직종이다. 직종 선정 후 2단계에서 기업을 알선하도록 한다.

나. 기업선택

2단계에서 기업을 선정할 때는 특성 간 가중치와 특성 내 일치도를 계산하여 개별 구직자별로 가장 적합한 기업을 선정하도록 하였다. 소프트웨어의 연관성 계산을 위해서는 구직 시 고려하는 여러 특성 간 중요도를 비교하는 가중치(= global weights)와 특성 내 일치도(= local consistency index), 그리고 일치도의 확률분포(= consistency probability distribution)를 계산하여야 한다. 특성 간 가중치란, 예를 들어 임금이나 경력 중에서 어느 특성이 얼마나 더 중요한지를 상대적으로 측정하는 것이다. 특성 내 일치도란 구직자의 특성(또는 희망)과 구직자의 희망(또는 특성)이 얼마나 일치하는가를 측정하는 것이다. 여기서 과거에 취업한 사람들의 특성 간 가중치와 특성 내 일치도, 그리고 일치도의 확률분포를 활용하게 되면, 새로운 구직자가 특정기업에 매칭될 가능성을 계산할 수 있다.

그런데 워크넷과 고용보험 DB의 자료만으로는 특성 간 가중치를 추정하기가 어렵다. 따라서 설문조사 결과를 이용하여 특성 간 가중치를 계산하였다.⁶⁾

- 5) MLP 신경망이란 입력층과 출력층 사이에 1개 이상의 은닉층을 갖는 다층구조로, 각 층에는 정보를 전달하는 인간의 신경과 같은 뉴런이 존재한다. MLP 신경망은 단층 신경망과 유사한 구조를 가지고 있지만 중간층과 각 단위의 입출력 특성을 비선형으로 함으로써 네트워크의 능력을 향상시켰다는 장점이 있다. 보다 자세한 설명을 위해서는 Haykin (1998)을 참조.
- 6) 설문조사 결과를 활용한 특성 간 가중치 계산은 지면 제약에 고려하여 본문에 수록하지

그리고 워크넷의 알선에 의하여 취업이 될 경우는 구직자의 희망 근로조건과 구인자의 희망 채용조건을 동시에 파악할 수 있다. 따라서 양 조건을 비교함으로써, 구직자 또는 구인자가 감내할 수 있는 조건의 차이에 대한 특성 내 일치도를 계산할 수 있다.

2단계의 잠재칭함수는 취업에 성공한 구직자(i)와 그 구직자가 취업한 기업(j)만을 분석대상으로 하여, 구직자 i 가 그 기업 j 에 취업할 확률을 구하는 함수이다. 이 함수를 활용한 취업확률이 실제취업 여부를 잘 예측하도록 잠재칭함수를 설계하는 것이 논문의 목표이다. 예측력이 높은 것으로 판정되면, 그 함수를 이용하여 구직상담을 하는 사람에게 적절한 기업을 선택하는 데 도움을 줄 수 있을 것이다. 따라서 i 번째 구직자와 그 구직자가 취업한 j 번째 기업에 대한 2단계의 잠재칭함수 M 은 다음과 같이 설정한다.

$$M(d_i, f_j) = \sum_{k=1}^l w_k \varphi_k(d_{ik} - f_{jk}), \quad i = 1, \dots, n, \quad j = 1, \dots, m, \quad (2)$$

여기서 d_{ik} 와 f_{jk} 는 각각 구직자 i 와 기업 j 의 k 번째 특성 값을 의미한다. 잠재칭함수는 특정 구직자가 특정 기업에 채용될 확률을 나타내는 것으로 0과 1 사이의 값을 가지게 된다. 따라서 주민이 n 명이고 기업이 m 개이므로 총 $n \times m$ 개의 함수 값이 계산된다.

특성의 개수 l 은 4개(학력, 경력, 희망임금, 근무지까지 거리)이며, 각 특성별로 서로 다른 값이 존재한다(예, 특성 : 학력, 경력 ..., 특성 값 : 학력 = 무학, 중졸이하, ..., 경력 = 경력 개월 수). 따라서 d_i 와 f_i 는 각각 주민과 기업의 i 번째 특성 값을 나타낸다. 특성 간 중요도를 파악하기 위하여 가중치(w_i ; global weights)를 포함한다. 이 값이 클수록 그 특성은 구직 혹은 구인 시 다른 특성보다 중요하다는 것을 의미한다. $\sum_i w_i = 1$ 의 관계가 성립한다.

그리고 $d_{ik} - f_{jk}$ 는 k 번째 특성에 대해서 i 번째 구직자의 특성(또는 희망)에서 j 번째 구인자의 희망(또는 특성)을 차감한 일치도를 의미한다. 이 값이 0이

않았다. 가중치에 대한 자세한 내용을 알려면 저자에게 문의.

면 양자가 일치하는 것이고 절대 값이 클수록 양자의 차이가 커지는 것을 의미한다. 여기서 $\varphi_k(d_{ik} - f_{jk})$ 는 k 번째 특성에 대해서 주민 i 와 기업 j 의 특성 값 차이에 대한 확률 값을 의미한다. 따라서 $\int_{-\infty}^{\infty} \varphi_k(x) dx = 1$ 의 관계가 성립한다. 워크넷과 고용보험 DB의 취업자 자료를 이용하여 일치도의 확률분포를 구하면, 개별 구직자별로 개별 기업에 대한 취업가능성을 계산할 수 있다.

구직자가 적합한 구인자를 선정하기 위해서는 구인자의 희망채용조건이 전부 밝혀져야 한다. 그러나 법에 의하여 성 또는 연령에 따라 차별할 수 없으므로 워크넷과 고용보험 DB정보는 활용이 불가능하다. 따라서 2단계에서 기업을 선정할 때는 학력, 경력, 희망임금, 근무지까지 거리의 정보만을 활용하여 취업 확률을 계산하도록 하였다. 그런데 근무지의 경우는 구직자의 주소나 근무예정지를 정확히 알 수 없는 문제가 있다. 따라서 이에 대한 대안으로 구직자가 거주하는 동의 주민자치센터에서부터 근무예정지 동의 주민자치센터까지의 거리를 구하여 대용변수(proxy variable)로 사용하였다.

2. 중분류 직종 선택

중분류 직종(1~23)의 로지스틱(logistic) 회귀분석 표본 수는 상용직 구직자 전체인 94,359명이다. 중분류 직종별 로지스틱 분석결과는 대부분의 직종에서 만족스러운 것으로 평가된다. R^2 는 0.08에서 0.34 사이에 위치하는 것을 알 수 있다.⁷⁾ 일반적으로는 표본 수에 비하여 해당 직종의 취업자 수가 적을 때 R^2 가 낮게 나오게 된다. 그럼에도 불구하고 일부 직종에서 R^2 가 0.30을 초과하여 예상보다 훨씬 예측력이 높은 것으로 판단된다.

해당 직종에 취업할 확률이 높은 구직자를 순서대로 배열하였을 때 이들이 해당 직종에 실제로 취업하여 있는지를 보여주는 이익도표⁸⁾ <표 3>에 의하면,

7) 중분류 직종 1~23에 대한 회귀분석결과의 세부 내용에 대해서는 저자에게 문의.

8) 이익도표는 예측확률이 높은 상위 10%, 20%, ... 등의 순으로 구간을 나누어 x 축에 배열하였을 때, 실제 취업자 중에서 해당 구간에 속하는 실제 취업자의 비중(Gain %)을 y 축에 그린 도표이다. 따라서 x 값이 작을수록 y 값이 크게 나타날 때, 예측확률은 실재를 잘 예측하는 것이다.

〈표 3〉 중분류 직종별 취업예측확률 이익도표

	취업예측확률 상위 %									
	상위 10%	상위 20%	상위 30%	상위 40%	상위 50%	상위 60%	상위 70%	상위 80%	상위 90%	상위 100%
직종 1	25.1	19.1	13.0	10.5	9.7	7.0	6.0	4.5	2.8	2.3
직종 2	31.3	21.0	14.6	9.5	7.0	5.8	4.6	3.1	1.8	1.2
직종 3	60.9	14.1	9.4	5.2	4.5	2.0	2.5	1.0	0.5	0.0
직종 4	57.4	17.3	9.8	6.1	3.9	2.4	1.4	1.4	0.2	0.0
직종 5	62.5	13.8	7.5	8.8	1.3	5.0	1.3	0.0	0.0	0.0
직종 6	72.3	6.2	4.4	5.5	2.7	2.5	2.2	2.4	1.3	0.6
직종 7	71.1	11.1	6.2	5.4	3.2	1.3	0.5	0.6	0.5	0.2
직종 8	70.6	11.5	5.6	4.9	3.5	1.7	0.3	0.3	0.3	1.0
직종 9	58.1	12.6	10.3	9.0	6.2	1.2	0.7	0.7	0.9	0.3
직종10	39.8	13.1	10.6	8.7	8.2	6.6	7.0	3.7	1.5	0.8
직종11	64.9	18.3	6.3	3.3	2.1	1.4	1.1	1.4	0.5	0.6
직종12	50.0	9.3	7.9	5.7	5.4	7.1	4.3	6.4	1.8	2.1
직종13	62.1	8.7	7.3	4.6	4.9	4.2	3.5	1.8	1.6	1.2
직종14	43.1	14.4	10.0	7.6	5.8	5.5	4.8	3.9	2.5	2.5
직종15	55.5	17.6	11.1	8.5	4.2	1.0	0.7	0.5	0.7	0.2
직종16	51.7	16.9	12.7	7.7	5.3	1.2	1.9	1.7	0.4	0.5
직종17	44.1	19.3	13.7	7.5	3.7	3.1	3.1	1.9	3.1	0.6
직종18	50.8	11.5	9.8	7.0	6.1	6.1	2.5	2.0	2.0	2.0
직종19	56.9	14.0	10.4	5.5	3.0	2.7	2.5	2.7	2.0	0.2
직종20	65.7	13.0	7.7	5.3	3.5	2.7	1.5	0.6	0.0	0.0
직종21	46.4	8.7	6.5	5.8	10.9	8.7	3.6	5.1	2.9	1.4
직종22	23.4	17.2	16.1	11.1	9.6	7.0	5.4	4.2	3.5	2.5
직종23	51.3	20.2	10.4	7.0	3.3	2.9	1.8	1.6	1.4	0.2

주: 직종분류는 <표 1>의 주를 참고.

많은 직종에서 이들이 실제로 취업되어 있는 비중이 크기 때문에 취업 예측확률의 정확도가 높은 것을 알 수 있다.

취업예측확률 상위 10%의 구간에서 해당 직종 전체 취업자 중 취업자의 비중은 직종 6(보건의료관련직)에서 72.3%로 가장 높고, 그 다음이 직종 7(사회복지 및 종교관련직)과 직종 8(문화예술디자인방송관련직)에서 각각 71.1%과 70.6%이다. 그 밖에 60%를 상회하는 직종이 직종 20(정보통신관련직)의 65.7%, 직종 11(경비 및 청소관련직)의 64.9%, 직종 5(법률경찰소방교도관련직)의 62.5%, 직종 13(음식서비스관련직)의 62.1%, 직종 3(금융보험관련직)의 60.9%인 것으

로 나타났다. 전체적으로 50%를 상회하는 직종의 수가 16개로 나타나 전체 직종 수 23개의 69.6%에 이르고 있다.

그러나 취업예측확률 상위 10%의 구간에서 해당직종 전체 취업자 중 취업자의 비중이 직종 22(환경인쇄목재가공공예 및 생산단순직)에서는 23.4%, 직종 1(관리직)에서는 25.1%로 낮게 나와 예측력이 떨어지는 것을 알 수 있다. 그러나 그럼에도 불구하고 예측확률이 낮아질수록 실제 취업자의 비중이 감소하는 패턴은 동일하게 나타나고 있다.

부산지역의 1년간 워크넷과 고용보험을 연계한 자료에 의하면 구직자 94,359명에 대하여 총 923,444건의 알선이 이루어져서 일인당 알선건수는 9.8건인 것으로 나타나고 있다. 그리고 워크넷 알선으로 인하여 취업한 구직자의 수가 12,895명인 것을 알 수 있다. 즉 알선 성공률은 1.4%에 불과하다. 이로부터 취업자 일인당 알선건수를 계산해 보면, 71.6건인 것으로 나타났다.

그러나 본 연구에서 계산한 대로 구직자에게 직종을 추천할 경우, 알선 성공률은 매우 높아질 것으로 기대된다. 예를 들어, 취업자 비중이 가장 높은 직종 7에서 상위 10% 구간의 구직자는 8,078명이고 실제 취업자는 1,375명이다. 따라서 이 구간의 예측 정확도는 $17.0\%(=1,375/8,078 \times 100)$ 에 달한다. 즉 6명을 이 직종에 알선할 경우 대략 1명이 직종 7에 실제로 취업될 것으로 추정할 수 있다. 그리고 직종 15의 상위 10% 구간에서는 예측 정확도가 $12.5\%(=1,038/8,327 \times 100)$ 인 것으로 나타나, 8명을 알선할 경우 대략 1명이 직종 15에 실제로 취업될 것으로 추정할 수 있다. 물론 2단계에서 기업을 알선하는 단계가 남아 있다. 그러나 이러한 추정은 실제 자료를 근거로 이루어진 것이기 때문에, 본 연구의 직종 추천방식을 채택할 경우 실제로 취업으로 연결될 가능성이 기존의 방식보다는 훨씬 높아질 것으로 예상된다고 하겠다.

3. 세분류 직종 및 기업 선택

가. 직종선택

본 연구에서는 세분류 직종의 사례로서 직종 155(금형및공작기계조작원)를 선정하여 분석하였다. 직종 155를 선정한 이유는 첫째, 부산의 전략산업분야에

해당하는 직종으로 부산이 경쟁력을 갖추고 있는 분야이기 때문이다. 둘째, 워크넷(구직표, 구인표)과 고용보험 DB의 표본 내에 112명/개의 취업자가 존재하여 소표본(small sample)에 의한 자료 왜곡의 가능성이 상대적으로 적기 때문이다.⁹⁾

전 구직자를 대상으로 직종 155에 대하여 로지스틱 회귀분석을 한 결과는 <표 4>에 나타나 있다. R^2 는 0.221로 나왔으며, 모형은 유의한 것으로 판명되었다. 그 결과에 의하면, 남자가 여자보다 취업확률이 유의하게 더 높은 것으로 나타났다. 청년층에 비하여 연령이 많아질수록 계수 값이 음수로서 절대 값이 커지고 유의도도 높아지는 현상이 나타나고 있다. 고졸 학력자가 중졸 이하 학력자보다 유의하게 취업확률이 높은 것으로 나왔다. 그러나 전문대졸이나 대졸 이상의 경우에는 추정치가 유의하지 않았다. 따라서 고졸학력자의 취업확률이 가장 높은 것으로 보인다. 그리고 자격증 유무나 다과는 취업확률에 별 영향을 미치지 않는 것으로 생각된다. 경력은 많을수록 취업확률이 높아지는 것으로 나타났다. 추정치가 커지고 유의도도 높아지기 때문에, 경력이 없는 구직자보다는 경력이 많은 구직자가 취업에 더 유리하다는 것을 알 수 있다.

그리고 200만~300만 원 미만 월급을 희망하는 구직자의 채용확률이 5% 수준에서 유의하게 높은 것으로 나타났다. 더 적은 월급을 희망하거나 더 높은 월급을 희망하는 구직자의 경우는 취업확률이 더 낮아지는 것을 알 수 있다. 울산과 경남을 근무지로 희망하는 구직자의 취업확률이 높아지는 것으로 나타나, 이 직종의 취업자가 상대적으로 이 지역에서 근무할 확률이 높은 것으로 보인다. 또한 직종 155를 희망한 경우 동 직종에 채용될 확률이 매우 유의한

9) 제Ⅲ장의 분석에서 사용되는 구직표와 고용보험 DB 연계자료에서는 구직자가 94,359명이고 직종 155에 취업된 사람은 564명(전체의 0.6%)이다. 이 중에서 워크넷 알선에 의하여 취업된 구직자의 수는 101명이다. 따라서 잠재칭함수 2단계 분석에서 활용되는 워크넷(구직표, 구인표)과 고용보험 DB 연계자료에서도 표본 수는 101명이어야 한다. 그러나 표본 추출 시기의 차이 등으로 인하여 구직표, 고용보험, 구인표 DB 연계자료의 직종 155 취업자는 112명인 것으로 나타났다. 그 중에서 100명만이 제Ⅲ장의 구직표와 고용보험 DB 연계자료와 매칭되는 것으로 확인되었다. 표본 수를 가능한 한 많이 확보하기 위하여 세 분류의 직종선택에서는 564명을 표본으로 하였으며, 기업선택에서는 112명을 표본으로 하였다.

〈표 4〉 전 구직자 대상 직종 155 로지스틱 회귀분석 결과

변수		계수 값(B)	표준 편차	유의 확률	Exp(B)	95%신뢰구간	
						하한	상한
남자더미		2.258***	.203	.000	9.568	6.430	14.238
연령	30대	-.193*	.110	.080	.825	.665	1.023
	40대	-.386***	.141	.006	.680	.515	.896
	50대	-.698***	.179	.000	.498	.351	.706
	60대 이상	-2.163***	.433	.000	.115	.049	.269
학력	고졸	.486**	.198	.014	1.626	1.102	2.398
	전문대졸	.089	.221	.686	1.094	.708	1.688
	대졸 이상	-.360	.233	.122	.698	.442	1.101
자격증	1개	.087	.124	.483	1.091	.856	1.390
	2개 이상	.097	.124	.435	1.102	.864	1.404
경력	1년 미만	.445**	.194	.022	1.560	1.067	2.280
	3년 미만	.559***	.132	.000	1.748	1.351	2.264
	3년 이상	.567***	.112	.000	1.763	1.417	2.194
희망 월급	150만 미만	.161	.274	.556	1.175	.687	2.010
	200만 미만	.393	.272	.149	1.482	.869	2.527
	300만 미만	.542**	.276	.049	1.719	1.002	2.950
	300만 이상	.166	.350	.635	1.181	.595	2.342
희망 근무지	울산경남	.555***	.203	.006	1.742	1.169	2.594
	수도권	-1.033	.719	.150	.356	.087	1.455
	기타	.047	.600	.938	1.048	.323	3.397
	상관없음	-.171	.253	.500	.843	.514	1.385
희망직종155더미		2.658***	.096	.000	14.264	11.818	17.216
상수항		-7.853***	.358	.000	.000		

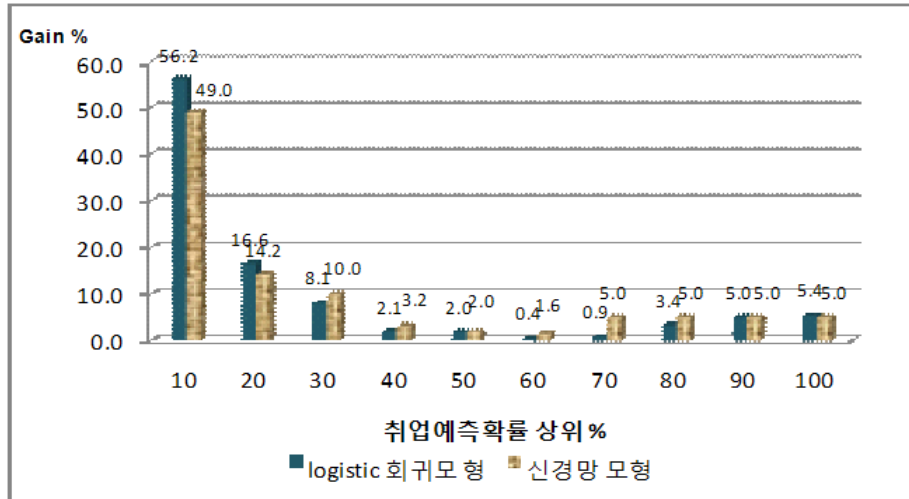
$$R^2=0.221 / \chi^2 = 1472.572$$

주: *, **, ***는 유의수준 0.10, 0.05, 0.01에서 각각 통계적으로 유의함을 의미함.

것으로 나타났다. 따라서 구직자가 직종 155를 희망할 경우, 그 희망을 존중할 필요가 있는 것으로 생각된다.

전 구직자를 대상으로 직종 155에 대한 이익도표를 그린 결과는 [그림 1]이다. 로지스틱 분석에 의하면, 직종 155 전체 취업자 중에서 56.2%가 상위 10%에 해당하는 것으로 나타났다. 그리고 상위 20%까지 포함하면 72.8%인 것으로 나타났다. 그러나 신경망 분석에 의하면, 그 비중이 각각 49.0%와 63.2%로 로지스틱 분석의 결과에 못 미치는 것을 알 수 있다.

(그림 1) 전 구직자 대상 직종 155 분석의 이익도표



나. 기업 선택

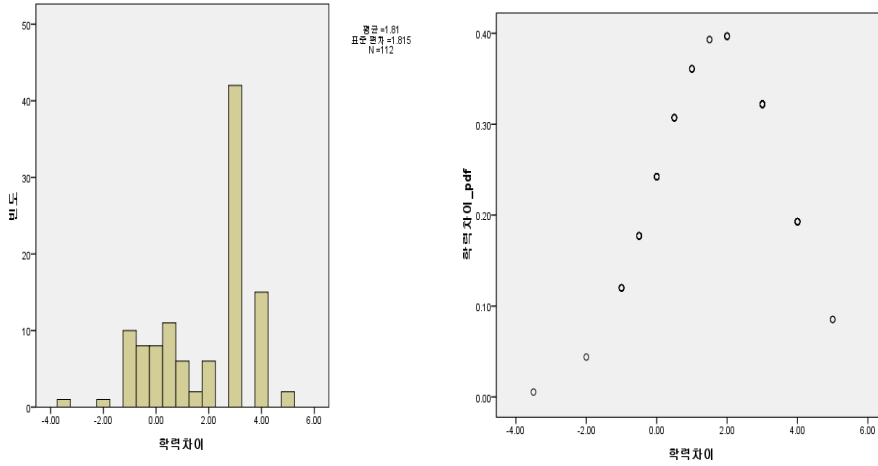
자료 이용이 가능한 직종 155의 취업자 112명과 기업 112개의 매칭표본(112개 표본)만을 대상으로, 2단계 잠재칭함수 일치도의 확률분포를 구하였다. i 번째 구직자와 j 번째 구인자의 k 번째 특성에 대한 확률분포가 바로 잠재칭함수의 일치도 확률분포인 $\varphi_k(d_{ik} - f_{jk})$ 이다.

1) 특성차이별 확률분포

학력의 일치도에 따른 빈도 및 표준화된 확률분포는 [그림 2]에 나타나 있다. 구직자 학력과 구인자 요구학력의 평균이 일치할 때는 0으로 나타난다. 그런데 학력차이가 0일 때의 빈도가 제일 큰 것이 아니라 학력차이가 3일 때의 빈도가 가장 큰 것으로 나타났다. 빈도차이가 불규칙적으로 나타난 것은 표본 수가 작기 때문이다.

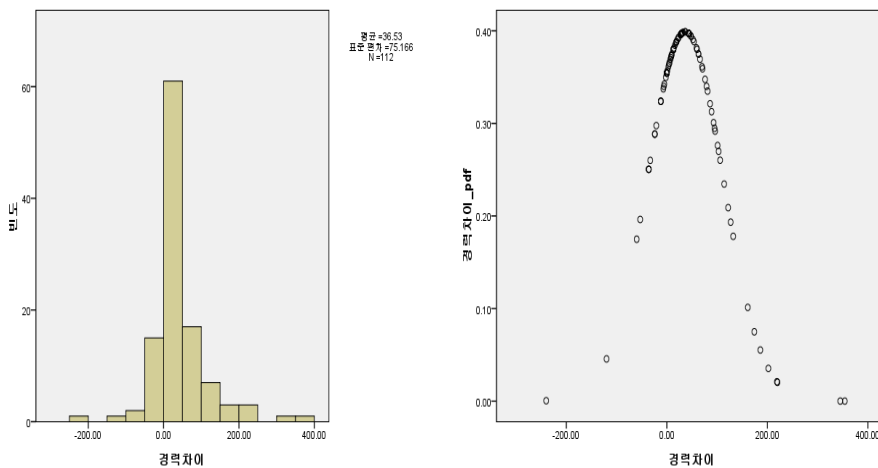
따라서 확률분포를 유연하게(smoothing) 하기 위하여 정규분포로 표준화시키도록 하였다. 표준화된 확률분포에서도 0에서가 아니라 2 이상에서 확률이 가장 큰 것으로 나타났다. 이러한 결과는 구직자가 구인자의 요구학력에 비하여 평균적으로 학력과잉(over-qualified) 상태에 있다는 것을 의미한다.

(그림 2) 학력차이의 빈도 및 표준화된 확률분포



경력의 일치도에 따른 빈도 및 표준화된 확률분포는 [그림 3]에 나타나 있다. 경력차이가 0에 근접한 수준에서 빈도가 가장 큰 것으로 나타나, 구직자 경력과 구인자 요구경력의 차이가 상대적으로 매우 작은 것을 알 수 있다. 확률분포가 정규분포와 유사하므로 확률 값을 구하기 위하여 정규분포로 표준화시키도록 하였다.

(그림 3) 경력차이의 빈도 및 표준화된 확률분포

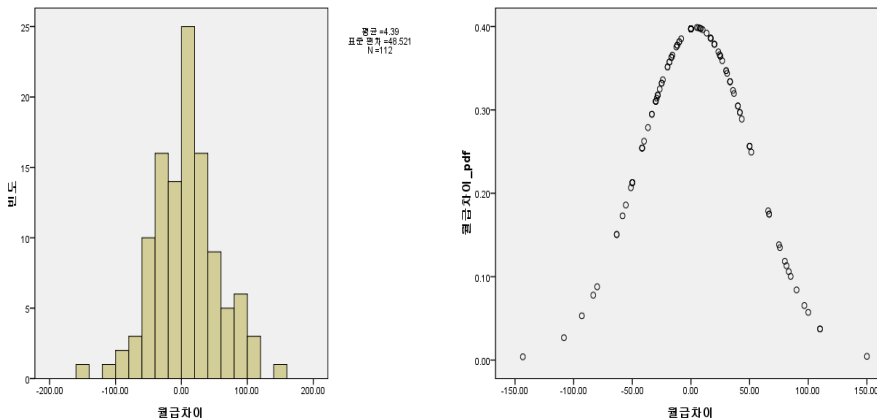


이러한 결과는 구직자의 경력이 구인자의 요구경력과 비교적 잘 일치한다는 것을 의미한다. 그 이유는 기업이 신입직원을 채용할 때는 경력직원이 지원을 하는 경우가 상대적으로 많지 않고, 기업이 경력직 사원을 채용할 때는 신입직원이 지원을 하는 경우가 많지 않기 때문인 것으로 보인다.

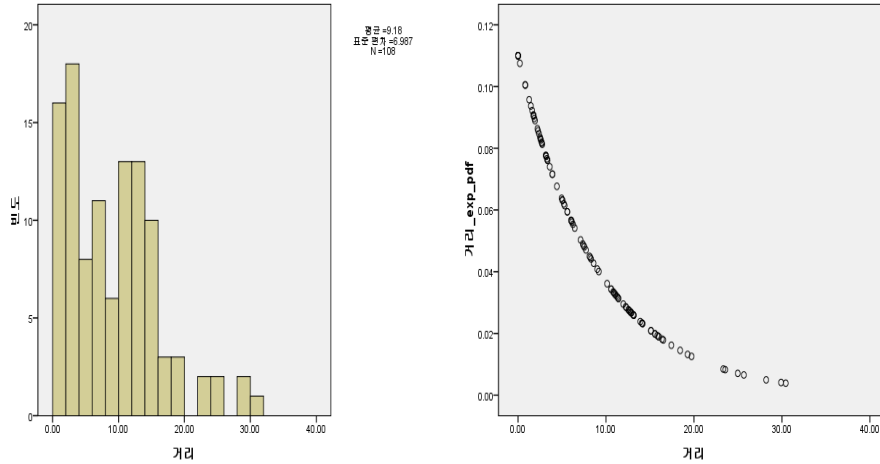
희망임금의 일치도에 따른 빈도 및 표준화된 확률분포는 [그림 4]에 나타나 있다. 구직자 희망임금과 구인자 제시임금 차이의 확률분포는 0이 평균인 정규분포에 매우 가까운 것으로 나타났다. 구직자가 희망하는 임금을 다 받지 못하고 취업할 것이라는 일반적인 인식을 근거로 판단한다면, 이러한 결과는 상당히 의외인 것으로 생각된다. 또한 비록 확률은 낮지만, 구직자가 희망하는 임금보다 훨씬 더 많은 임금을 제시하는 기업으로도 취업이 된다는 것은 더욱 의외인 것으로 생각된다.

이러한 결과는 취업시장에서 구직자의 협상력이 약하여 상대적으로 손해를 보면서 취업을 할 것이라는 주장에 대한 근거가 그렇게 강하지는 않을 수도 있다는 점을 의미한다. 그러나 이 결과를 해석하는 데 있어서 유의할 점은, 본 결과가 희망임금의 차이라는 점이다. 실제로 구직자와 구인자가 임금에 관하여 어떻게 합의를 본 것인지에 대해서는 확인할 수 있는 방법이 없다. 따라서 실제로 임금의 결정과정에서는 구직자보다는 구인자의 의사가 더 반영되었을 여지는 존재하는 것으로 보인다.

(그림 4) 희망임금차이의 빈도 및 표준화된 확률분포



(그림 5) 거리차이의 빈도 및 표준화된 확률분포



거리에 따른 빈도 및 표준화된 확률분포는 [그림 5]에 나타나 있다. 예상하는 바와 일치하게, 거리가 멀수록 취업확률 값은 작아지는 경향인 것으로 나타났다. 따라서 정규분포함수가 아닌 지수함수를 이용하여 확률분포를 구하였다. 부분적으로는 거리가 멀어져도 취업확률 값이 오히려 높게 나타난 이유는 표본 수가 작기 때문이다. 이 외에도 거주지 주변에 직종 155를 고용하는 제조업체가 많지 않기 때문인 것으로 판단된다.

2) 잡매칭함수

잡매칭함수를 계산하기 위해서는 특성 간 가중치(ω_k)가 필요하다. 지금까지 본 연구에서는 취업에 실제 성공한 구직자/구인자만의 자료를 활용하여 잡매칭함수를 도출하였기 때문에, 구직자와 구인자 모두에게 공통적으로 적용될 수 있는 것이라고 할 수 있다. 그러나 특성 간 가중치의 경우는 구직자의 가중치와 구인자의 가중치가 서로 다를 수 있다. <표 5>는 설문조사 결과에 의한 구직자와 구인자의 가중치이다. 본 연구에서는 두 가중치를 모두 사용하고 그 결과를 비교하였다.

개별 특성별 차이 값의 확률분포가 곧 잡매칭함수의 일치도의 확률분포, $\varphi_k(d_{ik} - f_{jk})$ 이다. 본 논문에서는 각 특성의 범주를 모두 10개로 분류하여 확

를 값을 산출하였다. 그 결과를 정리한 것이 <표 6>이다.

따라서 모든 구직자와 구인자 간의 잠재칭함수 값은 <표 5>와 <표 6>의 결과를 이용해서 식 (2)와 같이 계산할 수 있다.

<표 5> 잠재칭함수 값 : ω_k

	학력	경력	임금	거리
구직자	0.220	0.263	0.264	0.253
구인자	0.196	0.297	0.310	0.197

주: 구직자의 경력 가중치는 이용가능하지 않으므로 대신 희망직업과의 일치도를 대응 변수로 사용하였음. 그리고 거리는 출근소요시간 가중치를 사용하였음.

<표 6> 일치도 확률분포 : $\varphi_k(d_{ik} - f_{jk})$

(단위: 개월, 만 원, km)

범주	학력차이			경력차이			월급차이			거리			
	최소값	최대값	확률	최소값	최대값	확률	최소값	최대값	확률	최소값	최대값	확률	
1	min	-3.504	0.00171	min	-241	0.00012	min	-143.4	0.00117	0.00	3.37	0.30809	
2		-3.503	0.00632		-240	0.00243		-143.3	-110.8	0.00766	3.38	6.75	0.21317
3		-2.559	-1.617	0.02152	-174	-109	0.02470	-110.7	-78.2	0.03568	6.76	10.13	0.1475
4		-1.616	-0.673	0.0562	-108	-43	0.12079	-78.1	-45.6	0.10734	10.14	13.51	0.10205
5		-0.672	0.271	0.11264	-42	23	0.28569	-45.5	-13.0	0.20886	13.52	16.90	0.07061
6		0.272	1.215	0.1733	24	89	0.32777	-12.9	19.6	0.26308	16.91	20.28	0.04886
7		1.216	2.159	0.20468	90	155	0.18248	19.7	52.2	0.2145	20.29	23.66	0.0338
8		2.160	3.102	0.18559	156	221	0.04921	52.3	84.8	0.1132	23.67	27.04	0.02339
9		3.103	4.046	0.12918	222	287	0.0064	84.9	117.4	0.03865	27.05	30.42	0.01619
10		4.047	max	0.10886	288	max	0.00041	117.5	max	0.00986	30.43	max	0.03634

주: 학력차이를 구할 때는 무학=0 또는 상관없음, 초졸=1, 중졸=2, 고졸=3, 전문대졸=4, 대졸=5, 석사=6, 박사=7의 값을 사용하였음. 구인자의 경우 최저학력과 최고학력의 평균을 사용하였음.

V. 결 론

본 연구에서 개발한 잡매칭함수는 현 상태에서도 나름대로 유용한 시사점을 제공해 주는 것으로 평가된다. 그러나 앞으로 워크넷에서 활용되기 위해서는 여러 방면에서 더 발전시켜야 할 것으로 보인다.

우선 앞에서 언급하였듯이, 1단계에서 직종을 선택하는 데 활용한 변수의 수가 제한적이라는 한계가 있다. 따라서 개인의 특성을 분리할 수 있는 변수를 추가하거나 혹은 기존 변수의 범주를 더 세분화하는 것을 고려할 필요가 있다.

2단계에서의 한계는 구직·구인에 중요한 역할을 할 것으로 예상되는 성과 연령의 정보를 활용할 수 없었다는 것이다. 제Ⅲ장의 직종별 취업자 분석이나 설문조사 결과를 볼 때, 성과 연령별 정보는 취업에 중요한 영향을 미칠 것으로 판단된다. 그럼에도 불구하고 이 정보가 알려지지 않기 때문에 잡매칭에는 상당한 장애가 발생하는 것으로 추측된다. 따라서 비공개를 전제로 워크넷에서 성과 연령에 대한 정보를 파악하여 잡매칭에만 활용하도록 하는 것이 필요하다.

본 연구에서는 연구범위를 중분류 직종에 대한 직종 선택과 하나의 세분류 직종선택 및 기업선택으로만 한정하였다. 그 이유는 본 논문이 잡매칭함수의 활용가능성을 시범적으로 보여주는 데 연구 초점을 두었기 때문이다. 그러나 잡매칭함수가 직업상담의 현장에서 효율적으로 활용되기 위해서는, 중분류만이 아니라 세분류 이하의 모든 직종에 대해서도 잡매칭함수를 개발하는 것이 필요하다고 할 수 있다. 잡매칭함수를 토대로 서론에서 언급한 연관성 지수를 계산할 때, 잡 미스매칭(job mismatch)의 문제는 해소되고, 구직자의 취업률은 크게 제고될 수 있을 것으로 기대된다.

참고문헌

- 김성안 외(2005). 「충청지역 취약계층 고용지원 방안에 관한 연구」. 한남대학교 사회과학연구소.
- 김형래 외(2009). 「취업률 제고를 위한 직업이동 행태분석 및 효과적 취업알선매칭 기법 선별」. 『이슈분석』 4. 한국고용정보원.
- 박성익·김종한·전현중·조장식·곽소희(2009). 「근로조건 유연성과 잡매칭함수 도입」. 『지역고용연구』 2(2). 한국지역고용학회.
- 박성익·전현중·조장식·곽소희(2009). 「고용촉진지구 중소기업·주민 잡매칭을 위한 실태조사 연구」. 경성대학교.
- 송호근·설동훈(2001). 「중소기업 인력부족실태 및 대응방안」. 노동부 연구보고서.
- 오성욱(2009). 「스웨덴의 강화된 인터넷 매칭 취업서비스 및 정책시사점」. 『직업과 고용서비스연구』 4(1): 39~48.
- 이장원(2001). 「근로취약계층의 노동연계복지 발전방안」. 한국노동연구원 연구보고서.
- 조인호(2009). 「취업정보 알선/매칭 효과성 제고를 위한 소프트매칭 및 직업 유사성 연구」. 미출판. 한국고용정보원.
- 최창곤(2006). 「전북지역노동시장의 구조와 특징에 대한 연구」. 『경제연구』 24(4): 67~98.
- 한국고용정보원(2009). 「소프트매칭 해외활용 사례조사 출장결과 보고서」. 한국고용정보원.
- Eyigungor B. (2010), “Specific Capital and Vintage Effects on the Dynamics of Unemployment and Vacancies.” *America Economic Review* 100: 1214~1237.

Haykin, Simon(1998). *Neural Networks: A Comprehensive Foundation* (2 ed.). Prentice Hall.

Mortensen, D. and C. Pissarides(1994). "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61 (3): 397~415.

Petrongolo, B. and C. Pissarides(2001). "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (2): 390~431.

Pissarides, C.(2000). *Equilibrium Unemployment Theory* (2nd ed.). MIT Press.

Shimer, R.(2005). "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review* 95 (1): 25~49.

abstract

Development of Job Matching Function

Sungik Park, Jangsik Cho, Hyunjoong Jun,
Jonghan Kim, Jangsoo Ryu

Currently worknet set up correlation values for soft matching arbitrarily. In order to compute rational correlation values, this study develops job matching function by combining worknet and employment insurance DB and deriving probability function of consistence between characteristics of job seekers and potential employers. In the first stage of job matching function development, medium - level job category is selected for job seekers, and in the second stage firm is selected. Major results are the followings : First, it is reliable to select job category for job seekers since logistic analysis offers credible results. Second, it is found that probability functions by characteristics are consistent with expectations and computed job matching function is appraised to have good fit of job matching process.

Keywords : soft matching, job matching function, worknet, employment insurance DB