

# 학과별 이중전공 현황 데이터의 연결성 분석

빅데이터전공 2020380613 강정민

## 1. 서론

고려대학교 세종캠퍼스의 졸업 요건은 제 2전공을 필수로 선택해야 한다. 제 2전공은 융합전공, 학생설계전공, 복수전공, 심화전공, 이중전공이 있다. 그 중 이중전공은 제1전공을 이수하면서 타 학과의 전공을 제2전공으로 동시에 이수하여 졸업시 2개의 학위를 수여하는 제도를 말한다.

여기서 학과별로 어느 과에 이중전공을 많이 신청하는지, 이에 대해 연결성을 분석해 보기로 하였다.

## 2. 본론

### 2.1 데이터 설명

고려대학교 디지털 정보처에는 이중전공 살펴보기 데이터가 기재되어 있다. 2020년 5월 자료이고, 서울캠퍼스와 세종캠퍼스의 학과 중 세종캠퍼스만 분류하여 데이터를 정리하였다.

16×16 행렬로 만들었고, 제 1전공과 제 2전공이 연결 되어있으면 1, 연결 되어있지 않으면 0으로 설정하였다. 또한, 세종캠퍼스 내에서 이중전공 신청 이력이 없는 학과는 삭제하였다.

### 2.2 분석 방향

데이터를 입력하고, 근접 중심성과 중개 중심성, 고유벡터 중심성을 분석하고 데이터의 밀도와 추이성, 상호성을 분석하기로 한다.

근접 중심성은 정의 1로서는  $n$ 개의 노드로 구성된 네트워크에서 노드  $i$ 의 중심성을 다음과 같이 정의한다.

$$C_{[1]}(i) = \frac{n-1}{\sum_{j \neq i} d(i,j)}, i = 1, \dots, n.$$

정의 2로는 최단거리의 평균을 내는 대신 그것의 역수의 평균으로 노드  $i$ 의 중심성을 정의하기도 한다.

$$C_{[2]}(i) = \frac{\sum_{j \neq i} 1/d(i,j)}{n-1}, i = 1, \dots, n.$$

중심 노드는 '중심성'의 정의에 따라 다를 수 있다.

중개 중심성의 정의는 다음과 같다.

$$C_B(v) = \sum_{i \neq v} \sum_{j \neq v, \neq i} \frac{g_{ivj}}{g_{ij}}, v = 1, \dots, n.$$

고유벡터 중심성은 중요한 노드에 연결된 노드가 중요하다는 관점에서 창안 되었고, 다음과 같다.

$$c_i \propto \sum_{j=1}^n a_{ij} c_j, i = 1, \dots, n.$$

밀도는 ‘가능한 총 연결선의 수’ 대비 ‘총 연결선의 수’로 정의되며, 추이성은 다음과 같이 정의 된다.

$a \neq c$ 에 대하여 if  $a \rightarrow b$  and if  $b \rightarrow c$ , then  $a \rightarrow c$ .

상호성은 다음과 같이 정의된다.

if  $a \rightarrow b$ , then  $b \rightarrow a$ .

## 2.3 분석 진행(With R)

```
###기말과제###

#학과 별 이중전공 연결성

library(sna)

major.data <- as.matrix(read.table("C:\\Users\\pinkk\\OneDrive\\바탕 화면\\jungm
major.data
rownames(major.data) <- c("경제통계학부 경제정책학전공","공공사회·통일외교학부 공
colnames(major.data) <- c("경제통계학부 경제정책학전공","공공사회·통일외교학부 공
gplot(major.data, displaylabels=T, vertex.cex=1, edge.col="pink", boxed.labels=F
major.data

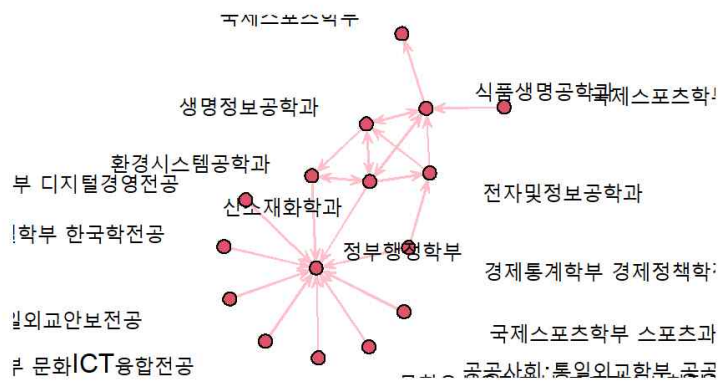
#근접 중심성
round(closeness(major.data, cmode="suminvdir"), 2) #나가는 방향
round(closeness(t(major.data), cmode="suminvdir"), 2) #들어오는 방향

#중개 중심성
round(betweenness(major.data), 1) #나가는 방향
round(betweenness(t(major.data)), 1) #들어오는 방향

#고유벡터 중심성
round(evcent(major.data), 2) #나가는 방향
round(evcent(t(major.data)), 2) #들어오는 방향

#major.data의 밀도
round(gden(major.data), 3)
#major.data의 추이성
round(gtrans(major.data), 3)
#major.data의 상호성
grecip(major.data, measure= "edgewise")
```

## 3. 결론



근접중심성은 신소재화학학과가 가장 높게 나왔으며, 국제스포츠학부와 정부행정학부가 가장 낮게 나왔다.

또한, 중개중심성은 나가는 방향 관점에서는 식품생명공학학과가 가장 높게 나왔으며, 들어오는 방향에서는 정부행정학부가 가장 낮게 나왔다.

고유벡터 중심성은 나가는 방향 관점에서는 신소재화학학과가 가장 높게 나왔으며, 들어오는 방향에서 또한 신소재화학학과가 가장 높게 나왔다.

밀도는 0.104, 추이성은 0.417, 상호성은 0.32로 나타났다.