

Graph Centrality를 통한 최적의 상권 입지 요건 분석

경희대학교 컴퓨터공학과
송근영(2017103997)

요약

서울특별시에서 제공하는 '100가지 생활밀접업종'과 '상권 분석 요인'을 Graph Centrality를 이용해 각 업종별 최적의 상권 입지를 도출하는 프로그램을 구축한다. 또한, 실제로 PC방과 숙박 시설을 표본 데이터를 입력하여 Graph Centrality를 분석한 결과, 해당 중심성에 따른 최적의 입지를 파악할 수 있었으며, 각 노드에 담겨있는 데이터를 파악함으로써 해당 상권이 지니는 특색 또한 분석할 수 있었다.

1. 서론

1.1 연구배경

현재 대한민국은 자영업 전성시대이다. 나이 불문할 것 없이 수많은 청년과 가정이 부푼 꿈을 안고 새로운 가능성에 투자해 새로운 자영업에 도전하고 있다. 중소벤처기업부에서 발표한 자료에 따르면 21년도 연간 자영업 창업 횟수 150만에 도달하며, 지난 해 대비 5.1% 증가했고, 이렇듯 자영업의 비율은 꾸준히 증가하고 있는 추세이다.

하지만, 이러한 많은 창업들이 모두 성공을 거두는 것은 아니다. 증가하는 자영업 창업의 비율 폭과 함께 연간 폐업자 횟수 현황 또한 90만에 육박하며, 이는 전체의 60퍼에 해당한다고 한다. 이렇듯 많은 자영업은 성공을 거두지 못 하고, 그대로 막대한 손실을 안은 채 사라지고 만다. 이러한 과정에서 수많은 자영업자들이 물질적 손해를 입으며, 다양한 사회경제적 문제들이 발생한다. 폐업이 반복될 경우 고용의 불안정을 야기하며, 기업 부채의 증가 등 사회적인 문제를 초래할 수 있다.

자영업에 실패하는 요인으로서는 다양한 요소를 꼽을 수 있다. 그 중 KB경영연구소의 설문 분석에 따르면 '상권 쇠퇴와 경쟁 과다 등 주변 환경 악화'를 1순위(45.6%)로 뽑고 있다. 이렇듯 자영업의 가장 큰 실패 요인 중 하나는 '상권 분석 실패'이다. 입점을 준비함에 있어, 들어서는 가게의 상권적 요인은 매우 중요하게 작용한다. 유사 업종이 해당 상권에 이미 얼마나 위치해있는지, 고객의 주 소비층은 어디에 해당하는지, 주변에 고객이 모일 수 있는 여러 편의시설이 존재하는지 등 여러 요인이 작용한다. 따라서, 해당 상권을 정확하게 분석하여 사업의 흥망 여부를 예측하는 것은 매우 중요하다.

1.2 연구목표

분석하고자 하는 업종을 입력할 시, Graph Centrality를 이용해 해당 업종이 입점하기 좋은 최적의 상권을 분석해준다. 이는, 상권 분석에 이용되는 여러 요인들을 이용하여 Graph Centrality에 도입해 다양한 중심성을 연구한다. 또한, Graph의 특성과 Gephi를 이용한 편리한 시각화와 데이터 분류로 해당 지역의 업종이 지니는 특성을 손쉽게 파악해, 자신의 니즈에 유리한 상권을 파악할 수 있도록

한다.

2. 관련연구

2.1 Graph

해당 연구에선 필요한 데이터를 노드와 엣지의 형태를 지닌 Graph로 변환하는 작업이 요구된다. 이를 손쉽게 수행해내기 위해서 Python의 라이브러리 NetworkX와 Pandas를 이용한다. 해당 라이브러리는 테이블과 시계열을 조작하기 위한 데이터 구조 와 연산을 제공하며, 그래프와 네트워크를 연구하기 위한 소프트웨어를 제공한다.

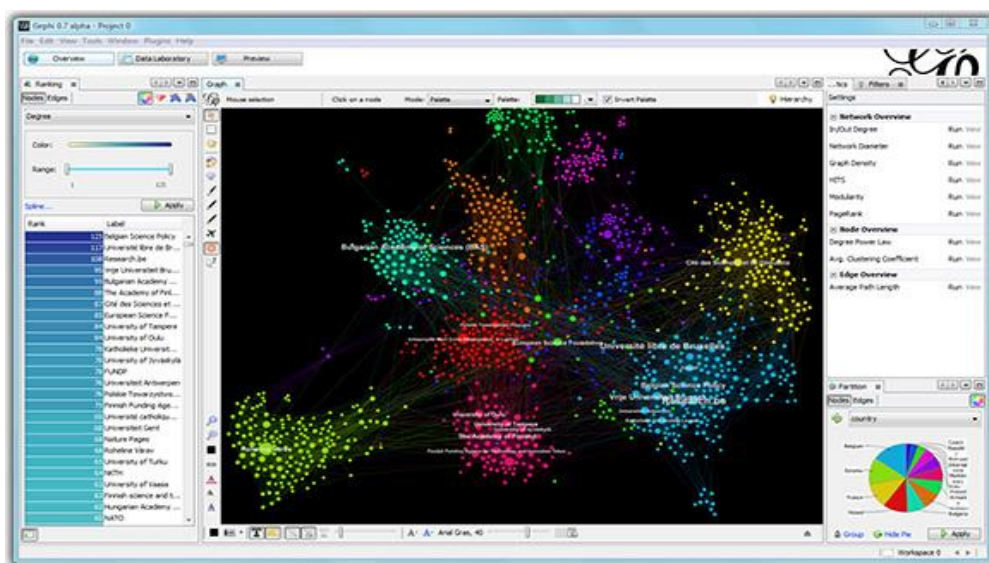
2.2 Graph Centrality

Graph Centrality란 그래프 혹은 사회 연결망에서 꼭짓점 혹은 노드의 상대적 중요성을 나타내는 척도이다. 이 중심성은 지수로 계산되는데 중점으로 두는 기준에 따라 다르다.

노드들 사이의 최단 경로를 가지고 계산하는 매개 중심성(Betweenness Centrality), 그래프의 노드와 다른 모든 노드들 사이의 최단 경로 길이 합의 역수를 이용해서 구하는 근접 중심성(Closeness Centrality), 한 노드에 연결된 모든 엣지의 개수로 중심성을 평가하는 연결 중심성(Degree Centrality), 주변 node의 중심성까지 고려하여 중심성을 선정하는 고유벡터 중심성(Eigenvector) 등 이 밖에도 다른 계산방법들이 다양하게 존재한다.[1]

2.3 Gephi

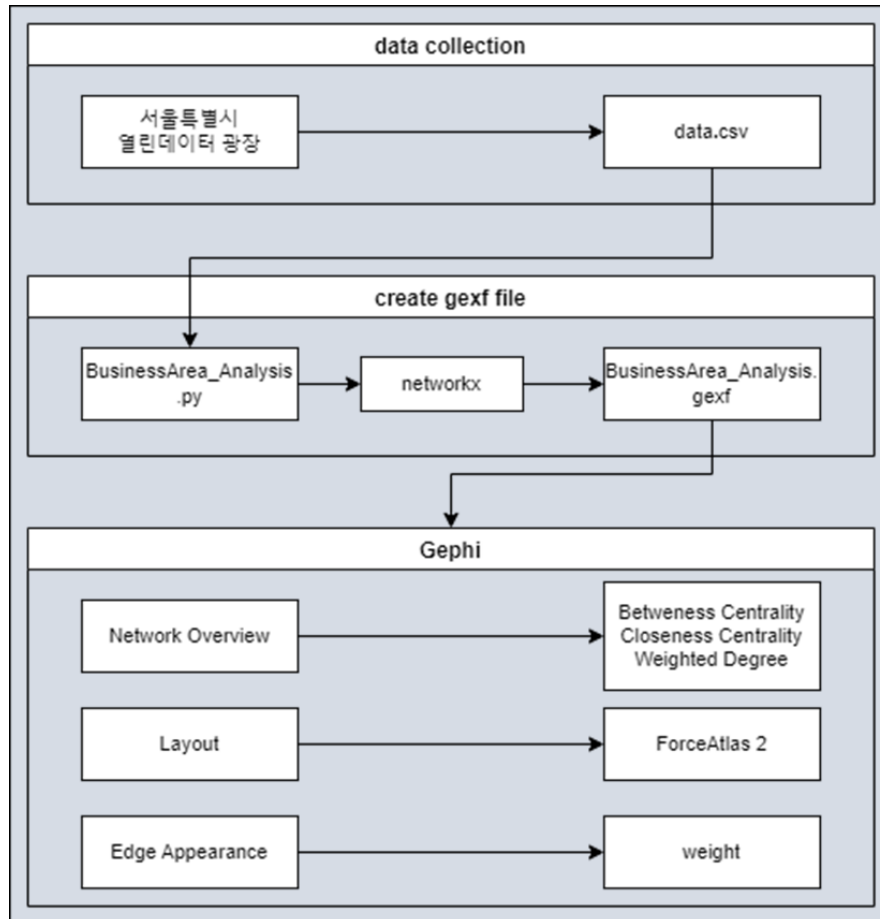
대규모 그래프를 시각화하고 분석하기 위한 오픈 소프트웨어이다.[2] 3D 렌더링 엔진을 사용해 실시간으로 그래프를 표시하고 탐색 속도를 향상해준다. 이를 사용해 각 행동 상권을 나타내는 그래프에 상세한 탐색과 분석이 가능하다.



[그림 1] Gephi 화면 예시

3. 프로젝트 내용

3.1 시스템 설계



[그림 2] 프로젝트 모듈 구조

본 연구의 전체적인 시스템 모듈 양상을 이와 같다. 서울 열린데이터광장에서 데이터를 수집한다. 이후 수집한 데이터를 그래프 시각화 프로그램 Gephi에 적합한 형태로 정제 및 가공하는 과정을 거친다. 이와 같은 과정을 거쳐 gexf파일이 생성되고, 이를 Gephi에 입력하여 다양한 중심성을 이용해 해당 상권을 분석해 원하는 결과를 도출한다.





3.2 데이터 추출 및 정제

데이터 제공 서비스로 서울 열린데이터 광장을 사용한다. 해당 사이트에선 환경, 교통, 인구 등의 데이터를 비롯한 서울시 시정활동 전반에 대한 정보를 제공한다. 데이터 사용에 문제가 발생하지 않도록 데이터에 대한 라이선스 정보를 게시하고, 집계하는 방식을 명백히 밝히고 있기 때문에 충분한 신뢰성을 지닌 데이터이다. 총 추출한 년도는 2017~2021년도 1~4분기의 데이터를 확보했으며, 표본 데이터로 2021년 1분기 데이터를 기준으로 정제 작업을 실행하였다.

	A	B	C	D	E	F	G	H	I	J
1	기준_년_코	기준_분기	상권_구분	상권_구분	상권_코드	상권_코드	서비스_업	서비스_업	분기당_매	분기당_매
2	2022	2 A	골목상권	2111090	강일동주민	CS300022	화장품	62192630	800	
3	2022	2 A	골목상권	2111090	강일동주민	CS300021	문구	30028758	3593	
4	2022	2 A	골목상권	2111090	강일동주민	CS300020	서적	1E+08	1819	
5	2022	2 A	골목상권	2111090	강일동주민	CS300018	의약품	5.38E+08	45236	
6	2022	2 A	골목상권	2111090	강일동주민	CS300016	안경	1.12E+08	955	
7	2022	2 A	골목상권	2111090	강일동주민	CS300010	반찬가게	1.72E+08	12687	
8	2022	2 A	골목상권	2111090	강일동주민	CS300007	육류판매	3.73E+08	12058	
9	2022	2 A	골목상권	2111090	강일동주민	CS300002	편의점	2.58E+08	40883	
10	2022	2 A	골목상권	2111090	강일동주민	CS300001	슈퍼마켓	1075164	413	
11	2022	2 A	골목상권	2111090	강일동주민	CS200031	세탁소	21783014	1105	
12	2022	2 A	골목상권	2111090	강일동주민	CS200028	미용실	3.74E+08	15109	
13	2022	2 A	골목상권	2111090	강일동주민	CS200024	스포츠클럽	10754899	37	
14	2022	2 A	골목상권	2111090	강일동주민	CS200019	PC방	1.96E+08	47097	
15	2022	2 A	골목상권	2111090	강일동주민	CS200016	당구장	40366846	2514	
16	2022	2 A	골목상권	2111090	강일동주민	CS200008	한의원	1.01E+08	7528	
17	2022	2 A	골목상권	2111090	강일동주민	CS200007	치과의원	5.18E+08	3872	
18	2022	2 A	골목상권	2111090	강일동주민	CS200006	일반의원	2.38E+08	18499	
19	2022	2 A	골목상권	2111090	강일동주민	CS200005	스포츠강	3.17E+08	1189	

[그림 3] 추출한 데이터의 일부

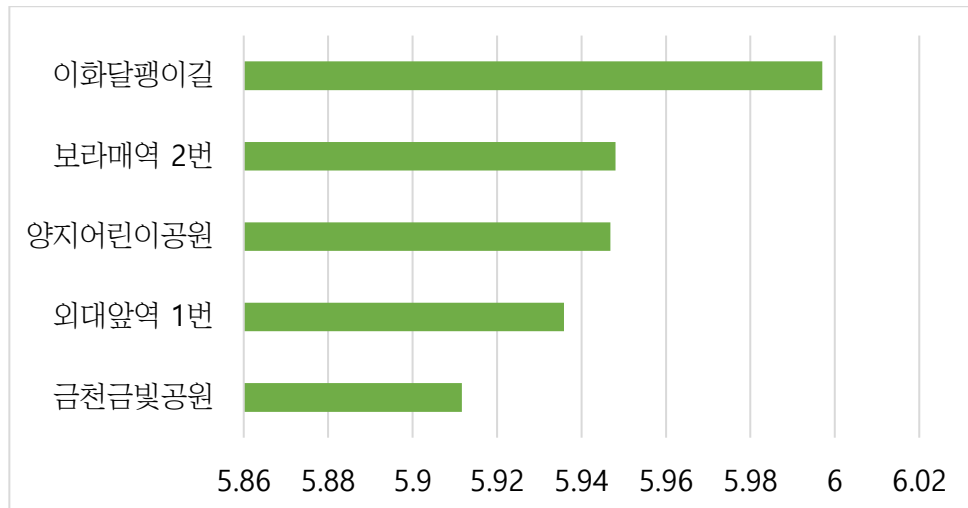
상권 분석을 위하여 총 7가지의 분석 요인을 추출해 csv파일로 저장했다. 총 수집한 데이터는 점포의 매출, 주변 집객시설 수, 유사업종 수, 인구소득, 직장인구, 생활인구, 상주인구 이렇게 구성돼있다. 또한, 분석하는 업종은 서울특별시에서 제공하는 ‘100개 생활밀접업종’ 데이터에 의거한다.

100대 생활밀접업종 선정현황			
대분류	100대 업종명(코드)	세세분류명(9차)	예시
	 한식음식점 (CS100001)	한식 음식점업 (56111)	· 설렁탕집 · 해물탕집 · 해장국집 · 보쌈집 · 냉면집
	 중식음식점 (CS100002)	중식 음식점업 (56112)	· 대중음식점(중식) · 중국음식점 · 탕수육전문점
	 일식음식점 (CS100003)	일식 음식점업 (56113)	· 초밥집(일식전문점) · 일식 횃집 · 로바다야끼 · 일식 우동집
	 양식음식점 (CS100004)	서양식 음식점업 (56114) 기타 외국식 음식점업 (56119)	· 레스토랑(서양식) · 패밀리 레스토랑 · 이탈리아 음식점 · 프랑스 음식점 · 베트남 음식점

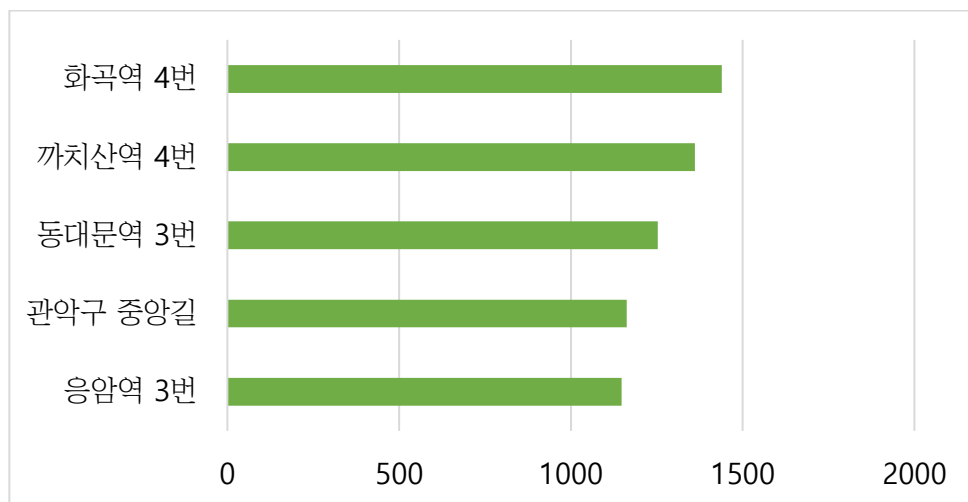
3.3 중심성 선정

GraphCentrality에는 또한 많은 기법이 존재한다. 매개 중심성(Betweenness

Centrality), 근접 중심성(Closeness Centrality), 가중 연결 중심성(WeightedDegree Centrality), 고유벡터 중심성(Eigenvector) 등이 있다. 각 그래프 사례에 적절한 중심성을 도입하는 것 또한 매우 중요하다. 어떤 중심성을 채택하나에 따라서 각기 다른 결과가 도출될 수 있다.



[표 1] PC방 상권 분석 가중연결 중심성(WeightedDegree Centrality) 상위 5곳

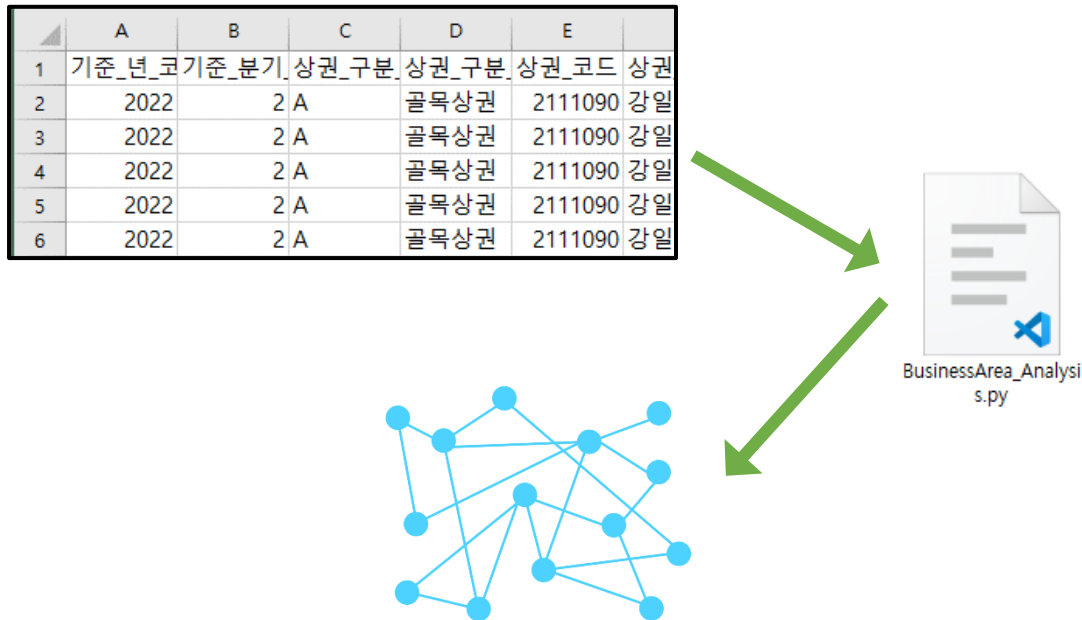


[표 2] PC방 상권 분석 매개 중심성(Betweenness Centrality) 상위 5곳

이와 같이 동일한 업종인 PC방에 대하여 중심성 분석을 진행하였지만, 서로 상이한 결과가 나온 것을 확인할 수 있다. 하지만, 해당 연구에선 각 엣지에 업종의 매출에 기반을 한 가중치 값을 부여하였다. 또한, 각 행정동 노드는 그 행정동 상권 분석 요인 노드와 연결되어있다. 따라서, 노드들 사이의 최단 경로를 가지고 계산하는 매개 중심성(Betweenness Centrality)은 행정동 노드와 행정동 노드 사이의 최단 경로를 이용해 중심성을 도출하므로 해당 연구의 목적과 부합하지 않으므로 적합하지 않은 중심성 기법임을 알 수 있다. 그러므로, 한 노드가 주변 노드와 연결된 노드의 수와 엣지의 가중치(매출)를 직접적으로 반영하여 중심성을 도출하는 가중연결 중심성(WeightedDegree Centrality)을 채택하여 해당 연구를 진행했다.

3.4 그래프 생성

앞서 데이터 추출 및 정제를 통해 점포의 매출, 주변 집객시설 수, 유사업종 수, 인구소득, 직장 인구, 생활인구, 상주인구의 정보를 포함한 csv파일을 생성했다. 이러한 데이터를 시각화하고 보다 정확하게 분석하기 위해 가장 적합한 자료구조 ‘그래프’를 채택했다. 상권 분석 요인을 그래프로 형성한다.



[그림 4] 그래프 생성 과정

파이썬을 이용해 호출된 csv파일에서 상권 분석에 필요한 여러 요소들을 가져온다. 이후, 행정동 상권 노드와 상권 분석 기준을 노드로 배정한다. 상권을 분석하는 기준은 유동인구, 거주인구, 직장인구, 유사 업종 점포 수, 소득수준, 집객시설 수, 매출이 있다. 다음, 각 행정동 상권 노드와 연결되는 상권 분석 노드를 연결하고, 각 엣지에 공식을 이용한 적절한 가중치를 부여한다. 이후, 이러한 정보를 포함한 gexf파일을 생성한다.

```
def Append_csvtoList(csvData, NodeList, Receive_List, RefNum):
    Receive_List = csvData.values.tolist()
    for i in range (len(NodeList)):
        for j in range(len(Receive_List)):
            if(NodeList[i][4] == Receive_List[j][4]):#상권코드가 같으면
                NodeList[i].append(Receive_List[j][RefNum])
                break
```

[그림 5] 노드 생성 코드 일부


```

for i in range(len(Receive_List)):
    for j in range(0,20):
        if( min+(max-min)/20*(j) <= Receive_List[i][num] and Receive_List[i][num] <= min+(max-min)/20*(j+1) and num != 12):
            if(num==9):
                GraphResult.add_edge(Receive_List[i][5] + " " + BusinessType, NodeName + str(j+1), weight = 1-Receive_List[i][8]*1.66 )
            elif(num==10):
                GraphResult.add_edge(Receive_List[i][5] + " " + BusinessType, NodeName + str(j+1), weight = 1-Receive_List[i][8]*1.66 )
            elif(num==11):
                GraphResult.add_edge(Receive_List[i][5] + " " + BusinessType, NodeName + str(j+1), weight = 1-Receive_List[i][8]*1.66 )
            elif(num==13):
                GraphResult.add_edge(Receive_List[i][5] + " " + BusinessType, NodeName + str(j+1), weight = 1-Receive_List[i][8]*1.66 )
            elif(num==14):
                GraphResult.add_edge(Receive_List[i][5] + " " + BusinessType, NodeName + str(j+1), weight = 1-Receive_List[i][8]*1.66 )

        elif( min+(max-min)/20*(j) <= Receive_List[i][num] and Receive_List[i][num] <= min+(max-min)/20*(j+1) and num == 12):
            GraphResult.add_edge(Receive_List[i][5] + " " + BusinessType, NodeName + str(j+1), weight = Receive_List[i][8]*1.66 )

```

[그림 6] 엣지 생성 코드 일부

또한, 해당 업종이 입점하기 좋은 최적의 상권을 분석하고, 각 상권의 특성을 손쉽게 파악하는 것이 목적이기 때문에 엣지에 각 점포의 평균 매출을 이용한 가중치를 부여해준다. 앞선 비교를 통해 선택된 가중연결 중심성(Weighted Degree Centrality) 가중치를 도출하는 과정은 이와 같다.

max = 해당 상권의 최고 매출
가중치 1 = 매출/(max+1)

월 소득 (가중치1) x A
유동인구 (가중치1) x B
직장인구 (가중치1) x C
거주인구 (가중치1) x D
집객시설 수 (가중치1) x E
유사업종 수 (1-가중치1) x F

A+B+C+D+E+F = 1

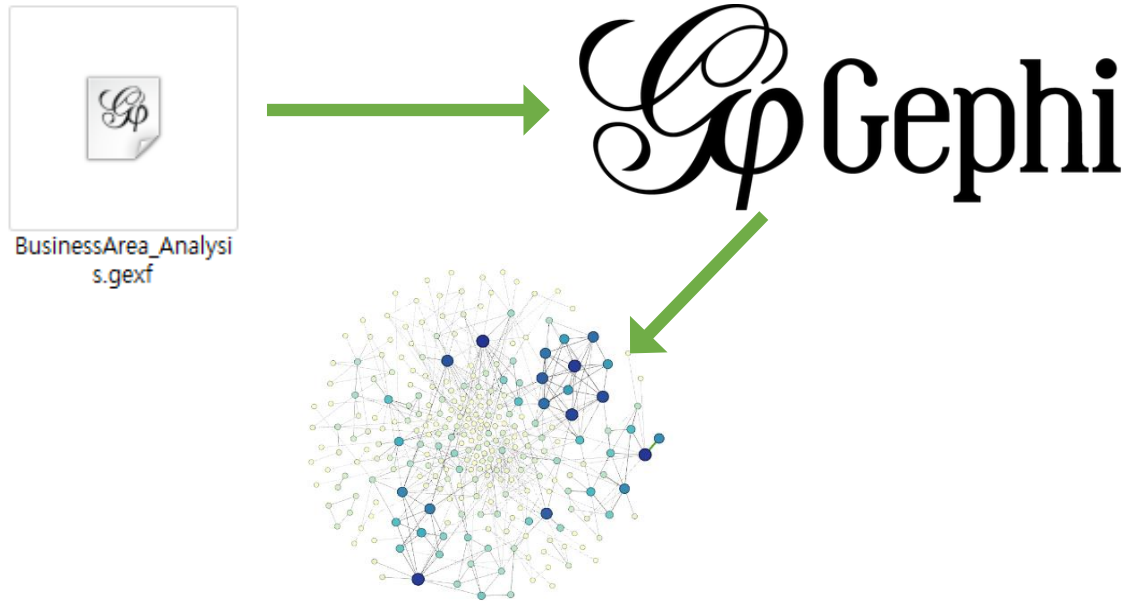
F	G	H	I
상권_코드	서비스_업	서비스_업	분기당_매출
이북5도청	CS100001	한식음식점	1245854151
이북5도청	CS100005	제과점	121486282
이북5도청	CS100008	분식전문점	195506473
이북5도청	CS100009	호프-간이	105816404
이북5도청	CS100010	커피-음료	30408910
이북5도청	CS200001	일반교습학	310913953
이북5도청	CS200003	예술학원	7468765
이북5도청	CS200006	일반의원	113081122
이북5도청	CS200007	치과의원	41264613
이북5도청	CS200008	한의원	95991627
이북5도청	CS200016	당구장	17092295
이북5도청	CS200017	골프연습장	11111915
이북5도청	CS200019	PC방	32207918

WeightedDegree Centrality 엣지의 가중치이다. 가중치1은 매출을 (해당 상권의 최고 매출 + 1)로 나눈 값이다. 이는 범위를 0~1로 제한하기 위해 위와 같은 과정을 거친다. 이후, 상권을 분석하는 총 여섯 가지의 상권 분석 기준이 존재한다. 각 가중치에는 해당 업종을 좀 더 유연하게 분석하기 위해 기본으로 존재하는 가중치에 사용자가 원하는 추가 가중치를 입힐 수 있다. 이 추가 가중치는 A, B, C, D, E, F이며, 이 추가 가중치의 총 합은 1로 일정하다. 또한, 유사업종의 수는 많을수록 상권 입지에 불리한 요소로 작용하기 때문에 유사업종의 경우 가중치를 (1-가중치1)로 설정한다.

이후, 해당 노드와 엣지의 가중치 정보를 포함한 리스트를 'Gephi'에 적합한 확장자인 '.gexf' 파일로 변형한다. 생성된 'WeightedDegree.gexf' 파일은 상권 행정동 노드, 상권 분석 노드, 가중치엣지를 포함한 데이터이다.

3.5 그래프 시각화

생성된 .gexf파일은 Gephi를 통해 시각화된다.



[그림 7] 그래프 시각화 과정

시각화의 큰 흐름은 이와 같으며, 자세한 과정은 아래와 같다.

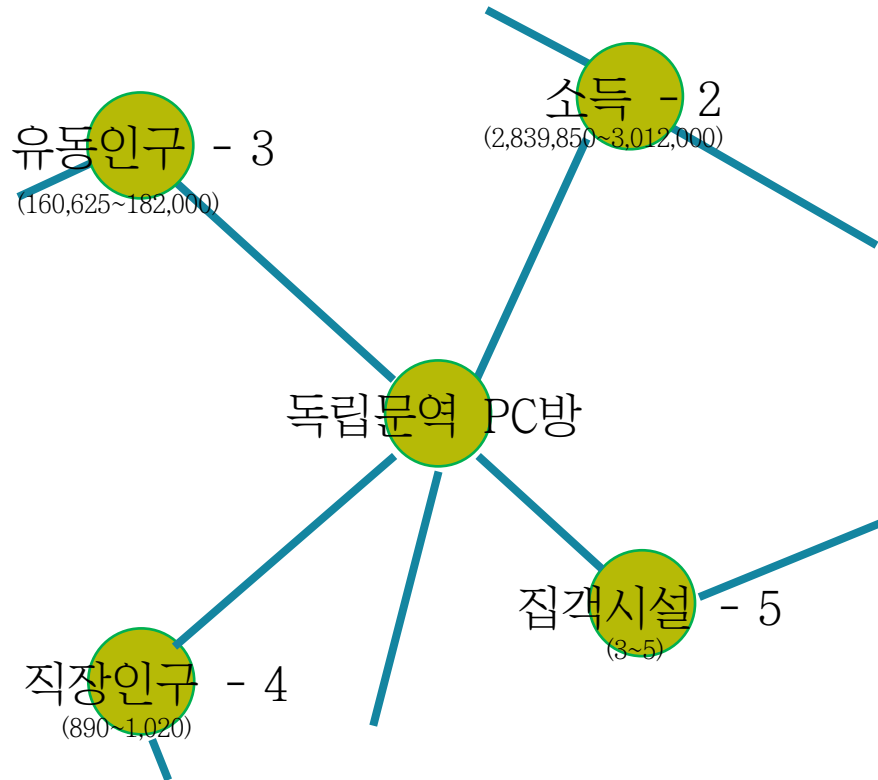
기준_년_코	기준_분기	상권_구분	상권_구분	상권_코드	상권_코드	서비스_업	서비스_업	분기당_매출
2021	1	A	골목상권	2110002	독립문역	CS200019	PC방	102354094

기준_년_코	기준_분기	상권_구분	상권_구분	상권_코드	상권_코드	서비스_업	서비스_업	점포_수	유사_업종
2021	1	A	골목상권	2110002	독립문역	CS200019	PC방	2	2

기준_년_코	기준_분기	상권_구분	상권_구분	상권_코드	상권_코드	총_생활인		
2021	1	A	골목상권	2110002	독립문역	1606205		

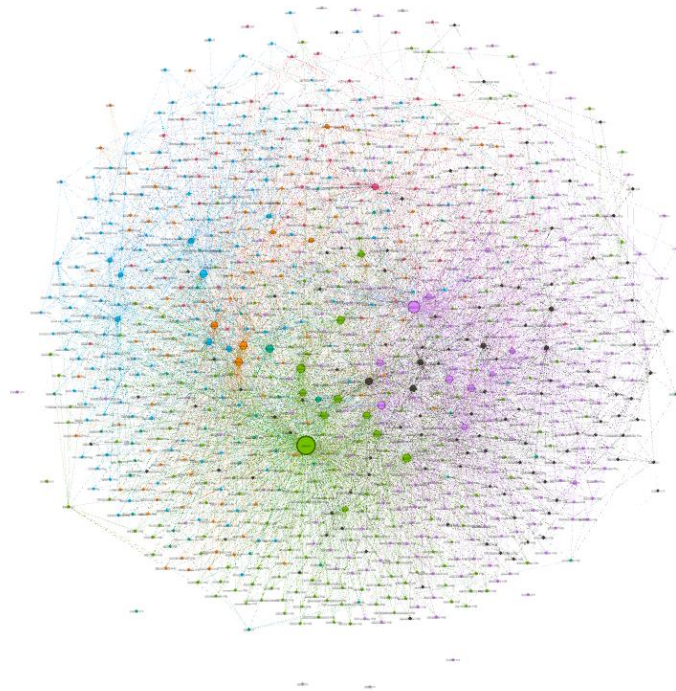
기준_년_코	기준_분기	상권_구분	상권_구분	상권_코드	상권_코드	월_평균_소		
2021	1	A	골목상권	2110002	독립문역	5115397		





[그림 8] 그래프 생성 예시

예를 들어, PC방이 키워드일 경우(검색키워드는 '서울특별시:100개 생활밀접업종'에 의거) 해당 프로그램에서 입력한 각 키워드에 부합하는 데이터를 추출해 위와같이 같이 노드와 엣지를 형성한 이후, 대응되는 노드끼리 연결해준다. 위는 PC방을 키워드로 했을 생성된 그래프의 예시이다. 노드는 928개, 엣지는 4865개가 형성되었다.



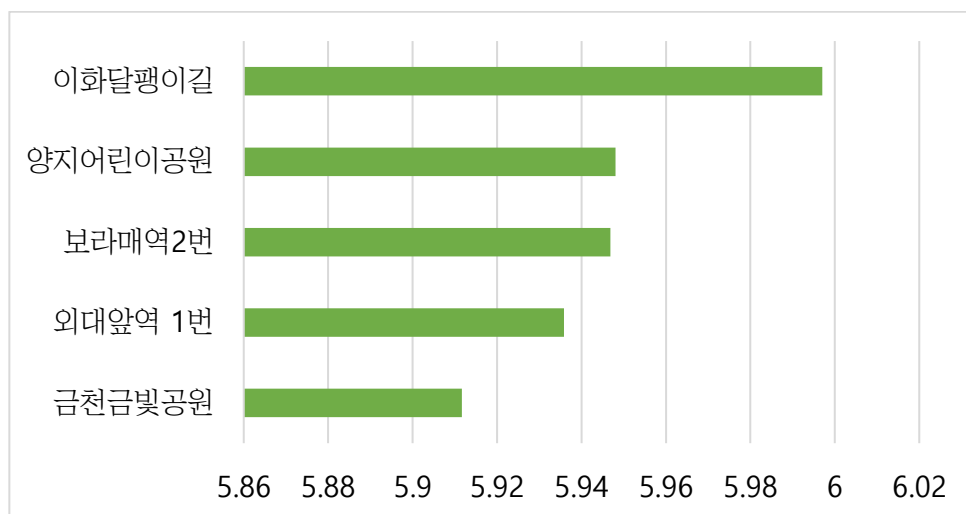
[그림 9] 그래프를 Gephi를 통해 시각화한 결과

앞서 분석했던 가중연결 중심성(Weighted Degree Centrality)을 보다 손쉽게 파악하기 위해 시각화하였다. 노드의 크기는 각 행정동의 상권 분석에 있어 중요도를 의미한다. 중요할수록 노드의 크기가 커지며, 매출을 이용한 엣지의 가중치가 높을수록 굵은 엣지가 나타난다.

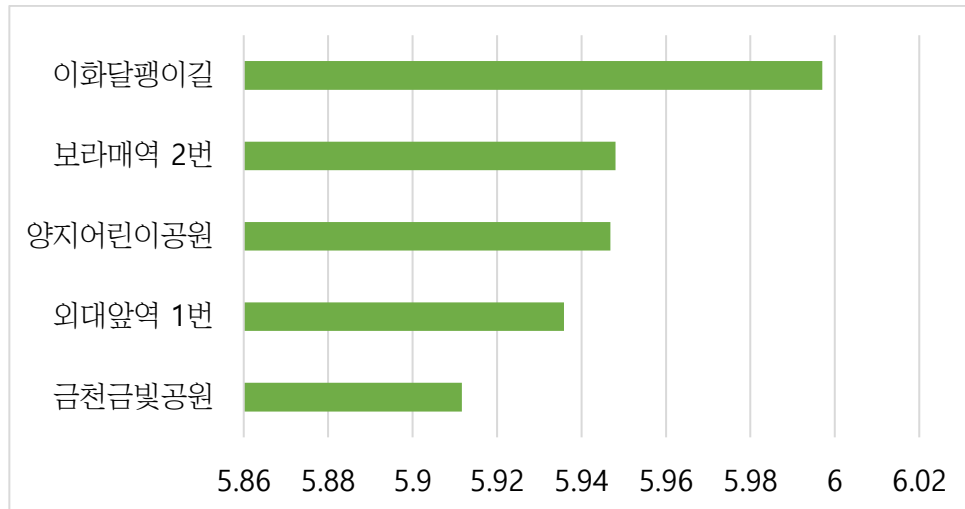
3.6 그래프 중심성 분석

총 100개의 업종 중 표본 데이터를 선정하여 입력해 그래프 중심성을 분석하였다. 해당 상권을 분석하고자 하는 사용자는 자신이 선택한 업종에 알맞게 가중치 퍼센티지를 선정하여 입력하게 된다. 따라서, 본인의 니즈에 알맞은 상권을 도출할 수 있다.

Id	Weighted Degree	Betweenness Centrality	Closeness Centrality	Harmonic
철소득16	5.991018	54.441265	0.319156	0.328747
보라매역 2번 PC방	5.935068	347.041258	0.341599	0.38364
양지어린이공원 ...	5.933519	347.274159	0.382988	0.423185
외대앞역 1번 PC방	5.919774	334.846241	0.370981	0.41199
금천금빛공원(은...	5.88962	355.907305	0.386192	0.42602
거주민구15	5.88339	64.84068	0.320486	0.329361
방학역 2번 PC방	5.87561	375.043187	0.371279	0.411755
NH농협은행 상도...	5.852788	371.353193	0.395798	0.433369
집객시설 수18	5.837389	101.087423	0.32386	0.331907
사가정역 1번 PC방	5.780598	370.534181	0.394108	0.432268
구로세무서 PC방	4.999995	200.501776	0.312881	0.349729



[표 3] PC방 상권 분석 가중연결 중심성(WeightedDegree Centrality) 상위 5곳

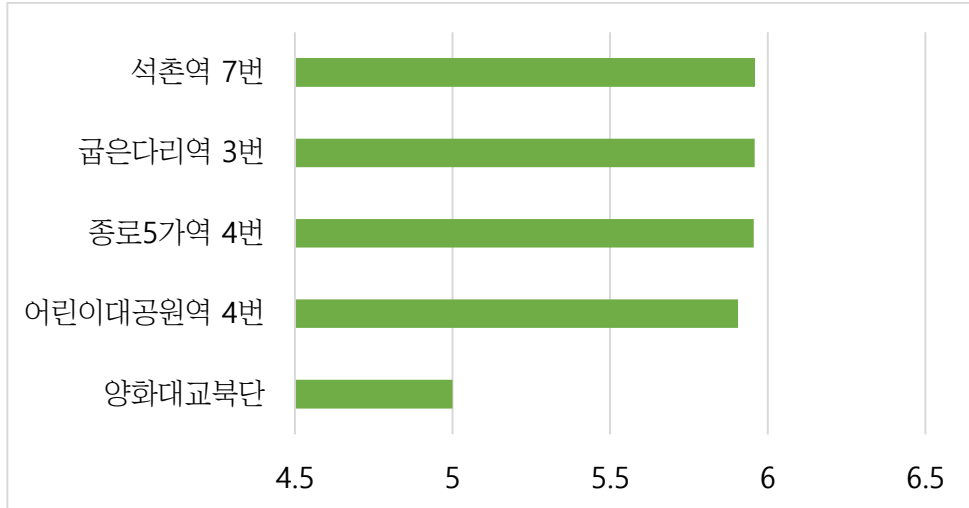


[표 4] PC방 상권 분석 가중연결 중심성(WeightedDegree Centrality) 상위 5곳

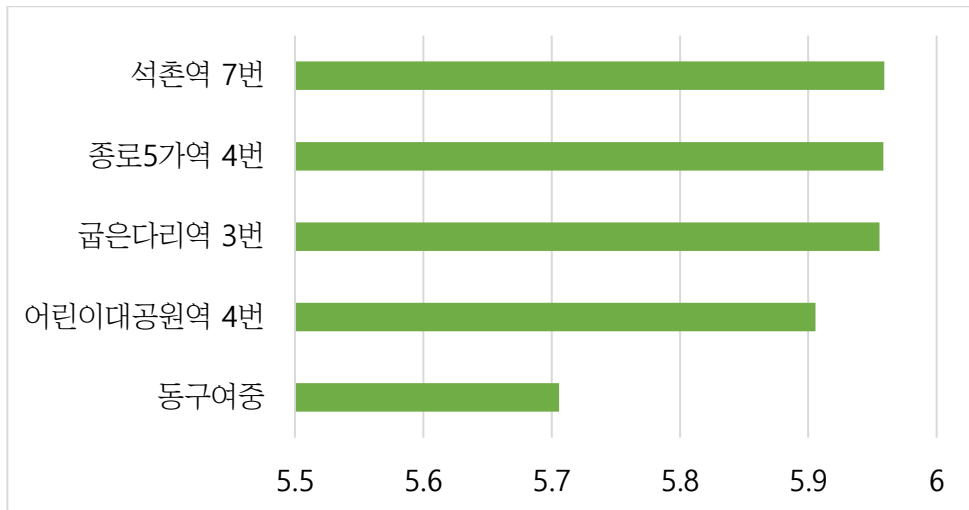
		표3	표4
유동인구	A	0.166	0.1
직장인구	B	0.166	0.1
거주인구	C	0.166	0.3
유사업종 수	D	0.166	0.2
집객시설 수	E	0.166	0.2
월 소득	F	0.166	0.1

먼저 PC방 상권을 가중연결 중심성(WeightedDegree Centrality)으로 분석한 결과이다. PC방은 해당 행정동의 상권 분석 요인인 유동인구나 직장인구보다 거주인구의 수요가 월 등히 많다. 따라서, 표3은 모두 동일한 가중치인 1.66을 부여하고, 개선된 가중치인 표4에서는 거주인구에 더욱더 높은 가중치를 부여한 모습이다. 이에 따라 동일한 가중치를 부여한 중심성 결과에서는 이화달팽이길, 양지어린이공원, 보라매역 2번, 외대앞역 1번, 금천금빛공원 순으로 중심성이 높게 도출되었지만, 직접 사용자가 조정한 가중치인 표4에서는 이화달팽이길, 보라매역 2번, 양지어린이공원, 외대앞역 1번, 금천금빛공원 순으로 가중치가 높게 나열된 것을 확인할 수 있다. 양지어린이공원과 보라매역 2번의 중심성 순위가 바뀌었는데 이는 보라매역 2번의 거주인구가 양지어린이공원보다 많고, 거주인구의 가중치를 높게 부여한 결과 이와 같은 중심성이 도출된 것을 확인할 수 있었다.

또한, 다음으로 모텔과 호텔 같은 숙박시설 업종을 가중연결 중심성(WeightedDegree Centrality)으로 분석한 결과이다



[표 5] 숙박시설 상권 분석 가중연결 중심성(WeightedDegree Centrality) 상위 5곳



[표 6] 숙박시설 상권 분석 가중연결 중심성(WeightedDegree Centrality) 상위 5곳

		표5	표6
유동인구	A	0.166	0.3
직장인구	B	0.166	0.2
거주인구	C	0.166	0.1
유사업종 수	D	0.166	0.2
집객시설 수	E	0.166	0.1
월 소득	F	0.166	0.1

숙박업종은 해당 행정동의 상권 분석 요인인 직장인구나 거주인구보다 유동인구의 수 요가 월등히 많다. 따라서, 표5는 모두 동일한 가중치인 1.66을 부여하고, 개선된 가중치인 표6에서는 유동인구에 더욱더 높은 가중치를 부여한 모습이다. 이에 따라 동일한 가중치를 부여한 중심성 결과에서는 석촌역 7번, 굽은다리역 3번, 종로5가역 4번, 어린이대공원역 4번, 양화대교북단 순으로 중심성이 높게 도출되었지만, 직접 사용자가 조정한 가중치인 표4에서는 석촌역 7번, 종로5가역 4번, 굽은다리역 3번, 어린이대공원역 4번, 동구여중 순으로 가중치가 높게 나열된 것을 확인할 수 있다. 마찬가지로 굽은다리역3번과 종로

5가역 4번의 중심성 순위가 바뀌었는데 이는 종로5가역의 유동인구가 양지어린이공원보다 많고, 유동인구의 가중치를 높게 부여한 결과 이와 같은 중심성이 도출된 것을 확인할 수 있었다.

이 외에도 사용자가 키워드와 가중치 퍼센티지를 다르게 입력하여 다양한 업종을 분석할 수 있다.

4. 결론 및 향후연구

4.1 결론

본 연구를 통하여, Graph Centrality를 통해 최적의 상권 입지 분석을 했다. 총 100가지 업종의 키워드를 입력 가능하며, 많은 데이터를 분석가능한 프로그램을 구축하였다. 해당 보고서의 표본 데이터로는 PC방과 숙박 업소를 분석하였다. 각자 좀 더 의미있게 작용하는 상권 분석 요인에 좀 더 높은 가중치를 부여하여 더 의미 있는 결과값을 도출 가능했다.

또한, 해당 중심성 순위를 이용해 높은 중심성을 차지한 상권의 특성을 손쉽게 파악할 수 있었다. 해당 유동, 거주, 직장 인구의 비율을 확인 가능하며, 집객 시설 및 경쟁 업체의 수도 파악할 수 있어 좀 더 전략적인 입점이 가능하다.

4.2 향후연구

해당 데이터사이트인 ‘서울 열린데이터 광장’에서는 좀 더 상세한 데이터를 제공하는데, 이를 포함해 GraphCentrality를 도출해 분석할 것이다. 해당 데이터에는 가벼운 인구 비율이 아닌 연령대, 성별에 따라 자세하게 분류가 되어있다. 이런 상세한 데이터 분류는 해당 상권을 파악하는데 더 큰 도움이 될 것이다.

또한, 해당 분석 프로그램에서는 사용자 자신이 원하는 업종의 키워드에 맞춰 알맞게 가중치를 입력할 수 있게 되어있지만, 문외한 업종일 경우 좀 더 전문적인 가중치 비율을 요구할 수 있다. 따라서, 보통 대외적으로 제시되는 이상적인 가중치 비율에 직접 학습하여 다가가는 알고리즘 또한 고안해 전문적인 지식이 없더라도 편하게 사용 가능하도록 범용성을 넓힐 생각이다.

5. 참고문헌

[1] 조태수, 한치근, 이상훈, “그래프 중심성들을 이용한 그래프 유사도 측정”, 한국컴퓨터정보학회논문지, 2018.

[2] <https://gephi.org/>