

# 과제

2022년 4월 19일 화요일    오후 12:50

2014097056 우성현

**3.5.8 빅데이터와 데이터 마이닝 기술은 온라인 소셜 네트워크에서 중요한 부분을 차지한다. 그 이유를 분산 시스템 기술이나 대규모 정보 시스템과 연관지어 설명하시오.**

▶ 빅데이터는 정형, 반정형, 비정형 데이터세트의 집적물로 기존 데이터보다 방대해 기존의 방법이나 도구로는 수집/저장/분석이 어려운 특징을 가지고 있다. 데이터 마이닝은 대용량의 데이터 속에서 유용한 정보를 발견하는 과정이다.

온라인 소셜 네트워크에서는 순간순간 엄청나게 많은 빅데이터가 쏟아져 나오는데, 이를 분석하기 위해서는 빅데이터와 데이터 마이닝 기술이 필요하다. 그런데 앞에서 언급했듯, 빅데이터는 기존의 컴퓨터에서는 다룰 수 없기 때문에, 여기에 필요한 기술이 분산 시스템과 대규모 정보 시스템이다.

분산 시스템은 다양한 유형의 시스템이 지역적으로 떨어져 있지만 네트워크로 연결한 시스템을 말하는데, 여기에는 분산 대용량 서버와 분산 대용량 저장소도 포함되어 있다. 이를 통해 대규모 정보시스템을 구성해 온라인 소셜 네트워크에서 생성되는 빅데이터를 다루고 분석할 수 있게 해준다.

**3.5.11 오프라인 소셜 네트워크와 온라인 소셜 네트워크를 개념적으로 구별 할 수 있는 가장 중요한 기준 혹은 속성 두 개를 설명하시오.**

▶ 온라인 소셜 네트워크는 오프라인 소셜 네트워크를 온라인으로 옮긴 것이기 때문에 프로필, 친구목록, 코멘트라는 중요한 세가지 요소를 가지고 있다.

또 다른 점은 온라인이기에 오프라인 소셜 네트워크보다 확장성이 뛰어나다. 개인을 중심으로 개인간 연결된 관계가 누적됨으로써 더 큰 네트워크를 형성한다.

**3.5.13 온라인 소셜 네트워킹이 무엇입니까?**

▶ 온라인 상의 인맥 구축을 도와주는 온라인 플랫폼을 말한다.

#### 5.4.5 아래 가중치가 있는 그래프에서 각 간선의 가중치는 시간을 나타낸다.

(1) 정점 A에서 정점 F까지 가는 경로를 모두 찾으시오.

P1 = (A, F)

P2 = (A, B, C, D, F)

P3 = (A, B, C, F)

P4 = (A, B, C, E, F)

P5 = (A, C, D, F)

P6 = (A, C, F)

P7 = (A, C, E, F)

(2) 정점 A에서 정점 F까지 가는 경로 중에서 반드시 정점 C를 거치는 경로를 모두 찾으시오.

▶ (1)번에서 P2, P3, P4, P5, P6, P7

(3) 정점 A에서 정점 F까지 가는 경로 중에서 시간이 가장 적게 드는 경로를 찾으시오. 또 시간이 가장 많이 소요되는 경로를 찾으시오.

▶ 시간이 가장 적게 드는 경로 : P5 = (A, C, D, F)

▶ 시간이 가장 많이 소요되는 경로 : P1 = (A, F), P4 = (A, B, C, E, F)

(4) 정점 A에서 정점 D까지 가는 경로 중에서 정점 C를 거치지 않는 경로를 모두 찾으시오.

▶ P1 = (A, F, D)

(5) 정점 A에서 정점 F까지 가는 경로 중에서 반드시 정점 C를 거치는 확률을 계산하시오.

▶ 6/7

(6) 정점 A에서 정점 F까지 가는 경로 중에서 반드시 정점 E를 거치지 않는 확률을 구하시오.

▶ 5/7

#### 5.4.13 아래 무방향 그래프에서 이심률, 지름, 반지름을 각각 구하시오.

▶ 이심률 : 2

▶ 지름 : 2

▶ 반지름 : 1

5.4.14 실습 5.4.13의 무방향 그래프에서 평균 경로 길이를 구하시오.

▶  $26/20 = 1.3$

	주석	규철	보라	은숙	경호	합	평균
주석	0	1	1	2	1	5	1
규철	1	0	1	2	1	5	1
보라	1	1	0	1	2	5	1
은숙	2	2	1	0	1	6	1.2
경호	1	1	2	1	0	5	1
합	5	5	5	6	5	26	-
평균	1	1	1	1.2	1	-	-