

과제: small world, motif, graph role, community detection을 활용하여, 네트워크 분석 및 시각화

이름: 정주호

학번: 2019312072

전공: 인공지능융합전공

## 1. 데이터 설명: 미국 대학 미식축구 (Schedule of Division I games for the 2000 season)

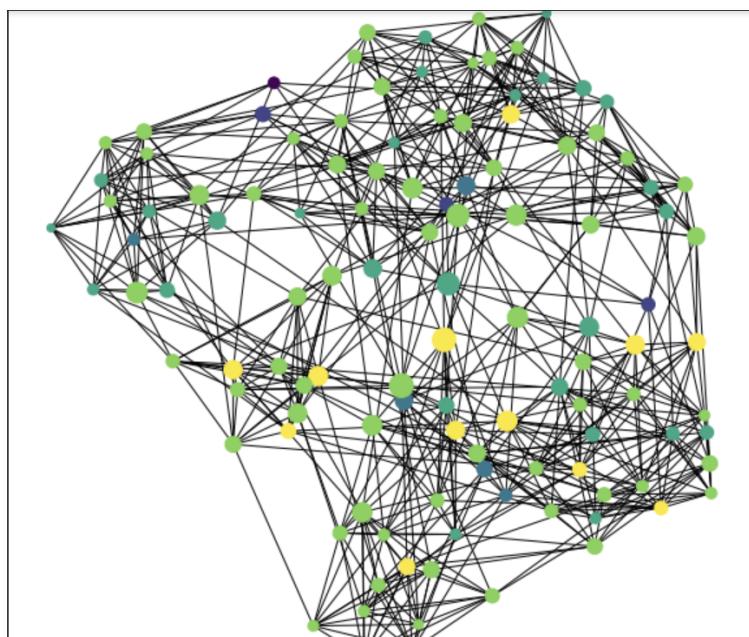
내가 살펴본 네트워크는 2000년 시즌의 Division I 게임 스케줄이다. 그래프의 node는 팀을 나타내고, edge는 연결된 두 팀 간의 정규 시즌 게임을 나타낸다.

## 2. 데이터 분석

각 팀은 약 8~12개 팀이 참가하는 컨퍼런스로 나뉜다. 네트워크 분석을 통해 node의 갯수의 평균을 구해보면 2000년 시즌에는 각팀들이 평균 약 7개의 컨퍼런스 내 게임과 4개의 인터컨퍼런스 게임을 하는 등, 같은 컨퍼런스 멤버들 간의 게임 빈도가 서로 다른 컨퍼런스보다 높다는 것을 알 수 있다.

다시 말해, 네트워크 분석을 통해 컨퍼런스 게임이 균일하게 분산되지 않는다는 것을 알 수 있다. 이는 미국 내 지리적으로 서로 가깝지만 서로 다른 컨퍼런스에 속한 팀들은 더 지리적으로 먼 거리로 분리된 팀들보다 서로 경기할 가능성이 더 높다는 결론을 도출할 수 있다.

\* Between Centrality



```

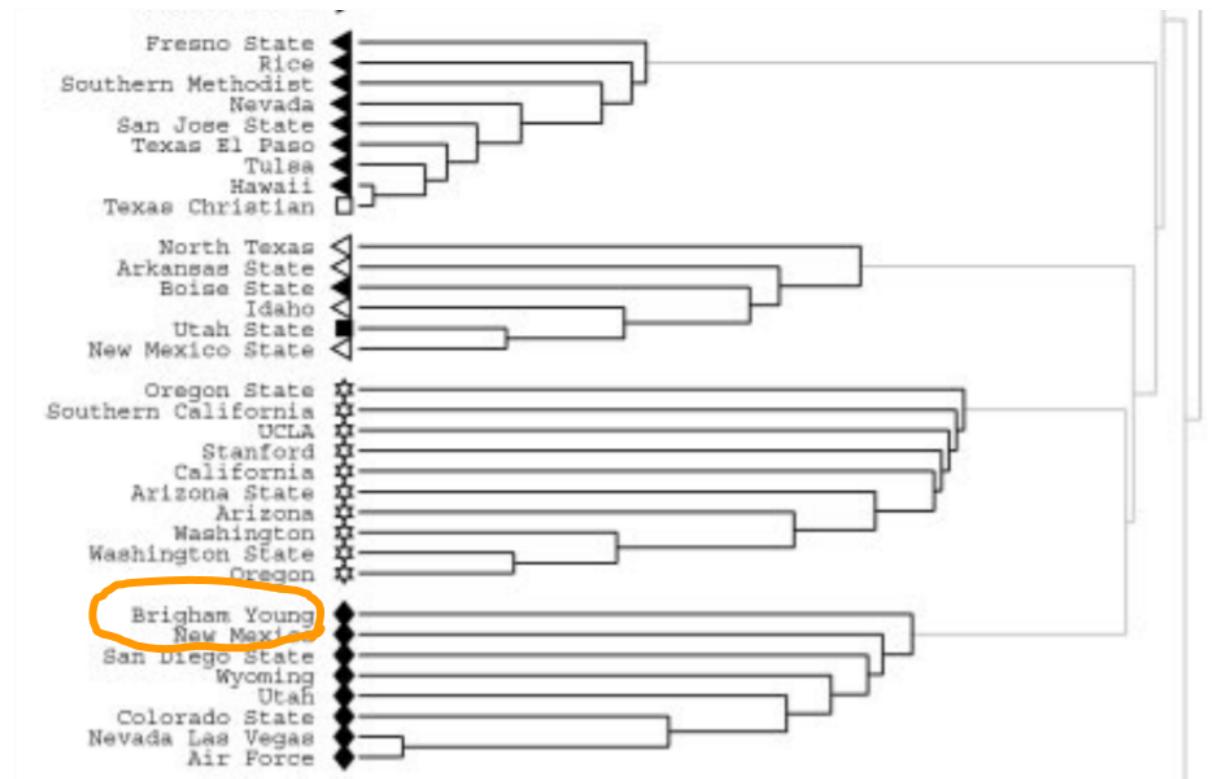
▶ betCent = nx.betweenness_centrality(g, normalized=True)
sorted(betCent.items(), key = (lambda x: x[1]), reverse=True)

⇨ ✓ [('NotreDame', 0.03353295672662947),
       ('BrighamYoung', 0.03248994918389482),
       ('Navy', 0.029161053264223705),
       ('LouisianaTech', 0.02882282342077095),
       ('CentralMichigan', 0.02518680197086883),
       ('NewMexicoState', 0.024139419776449275),
       ('Cincinnati', 0.02383635493910067),
       ('KansasState', 0.023070098856845606),
       ('Alabama', 0.02304632173060802),
       ('Wyoming', 0.022213447427560148),
       ('UCLA', 0.022126807176414888),
       ('Wisconsin', 0.020895265847707085),
       ('MiamiOhio', 0.020755436736059025),
       ('ArkansasState', 0.02070022463158684),
       ('Indiana', 0.020498159932875414),

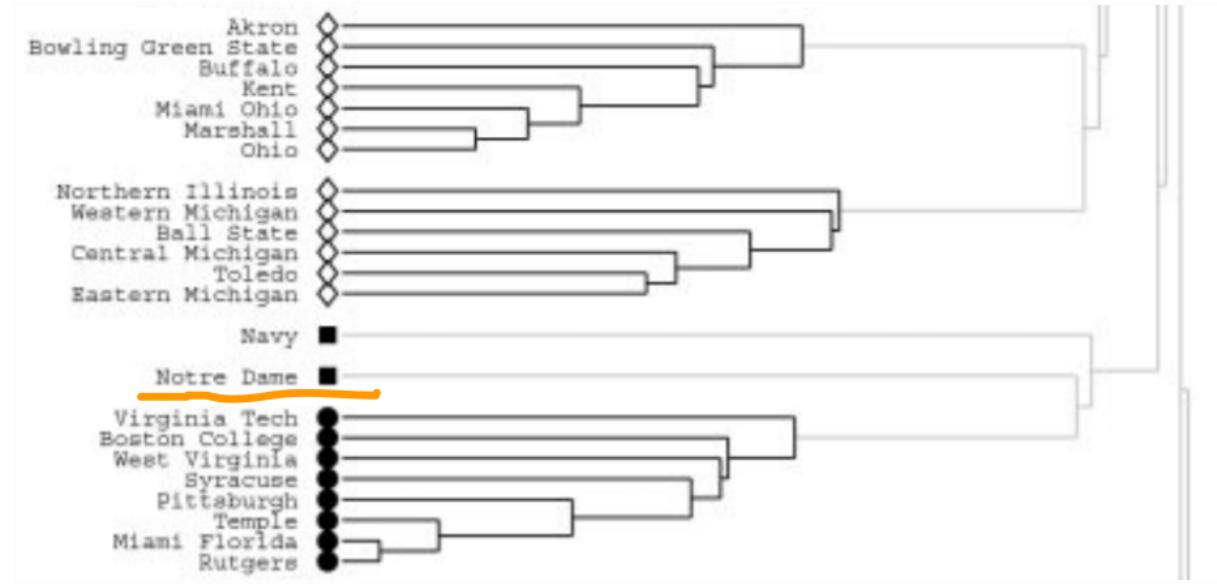
```

-> Betweenness Centrality : 노드 간의 경로에 있을 수록 큼.

현재 네트워크에서 Betweenness Centrality 가장 높은 팀이 ‘NotreDame’, ‘BighamYoung’인데 먼저 신기한 것은 ‘BighamYoung’은 node가 7개로 게임 수는 다른 팀들보다 적으나 아래의 사진처럼 본선에 진출하여 대부분 다른 컨퍼런스의 팀들과 게임하였기 때문에 컨퍼런스들 간의 연결에 중심에 위치하여 있다고 분석할 수 있다.

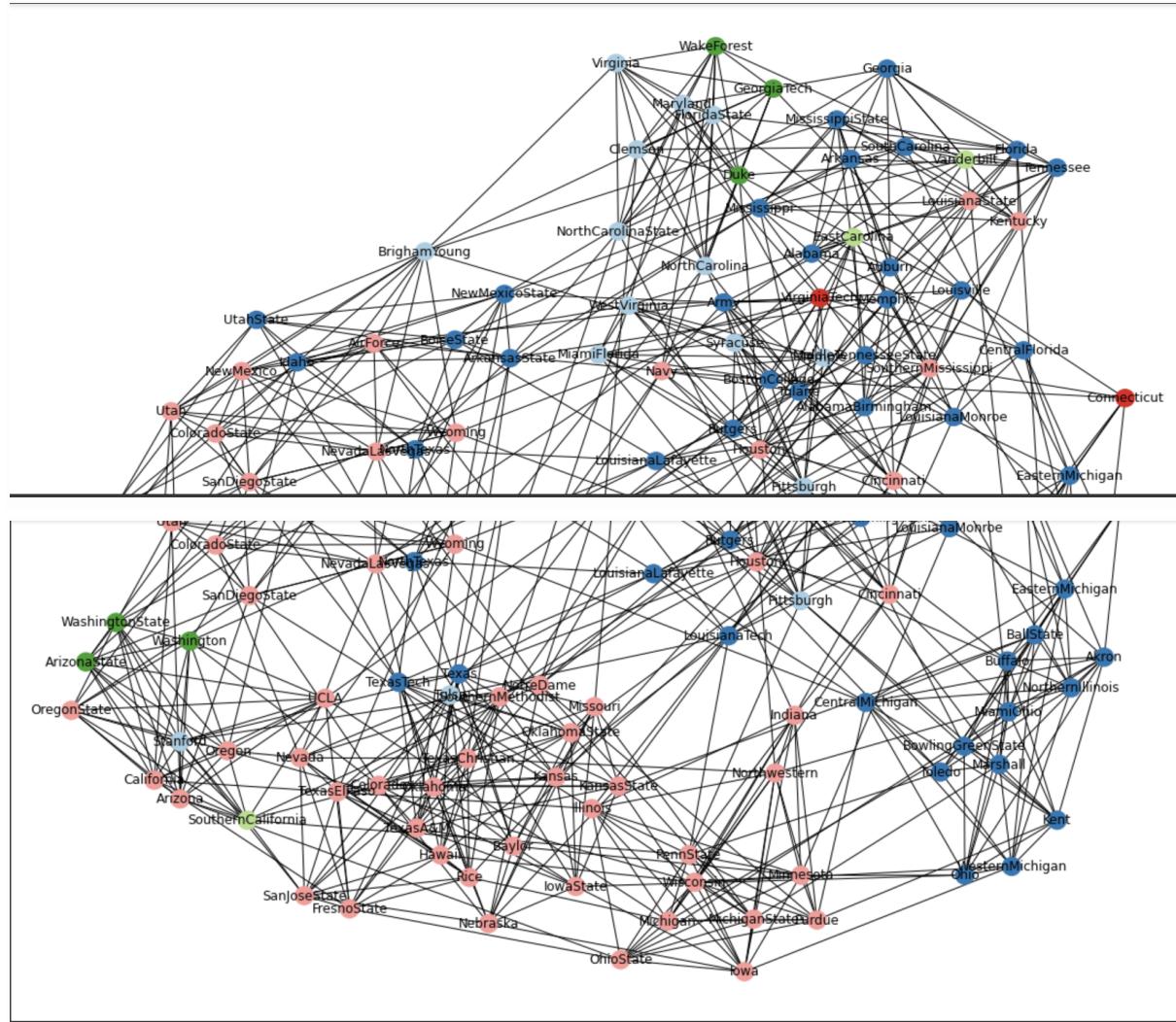


Betweenness Centrality 가장 높은 'Notre Dame'은 아래 그림처럼 부전승으로 올라가 결승까지 진출한 팀이다. 따라서 다른 컨퍼런스의 팀들과 경로에 위치하여 있다는 것을 알 수 있다. 여기서 가쁜 컨퍼런스와의 경로에 위치한다는 말의 의미는 본선에 진출한 팀은 각기 다른 컨퍼런스에서 올라온 것이기 때문에 본선에서 많이 게임을 한 팀이 경로에 위치할 확률이 높은 것이다.



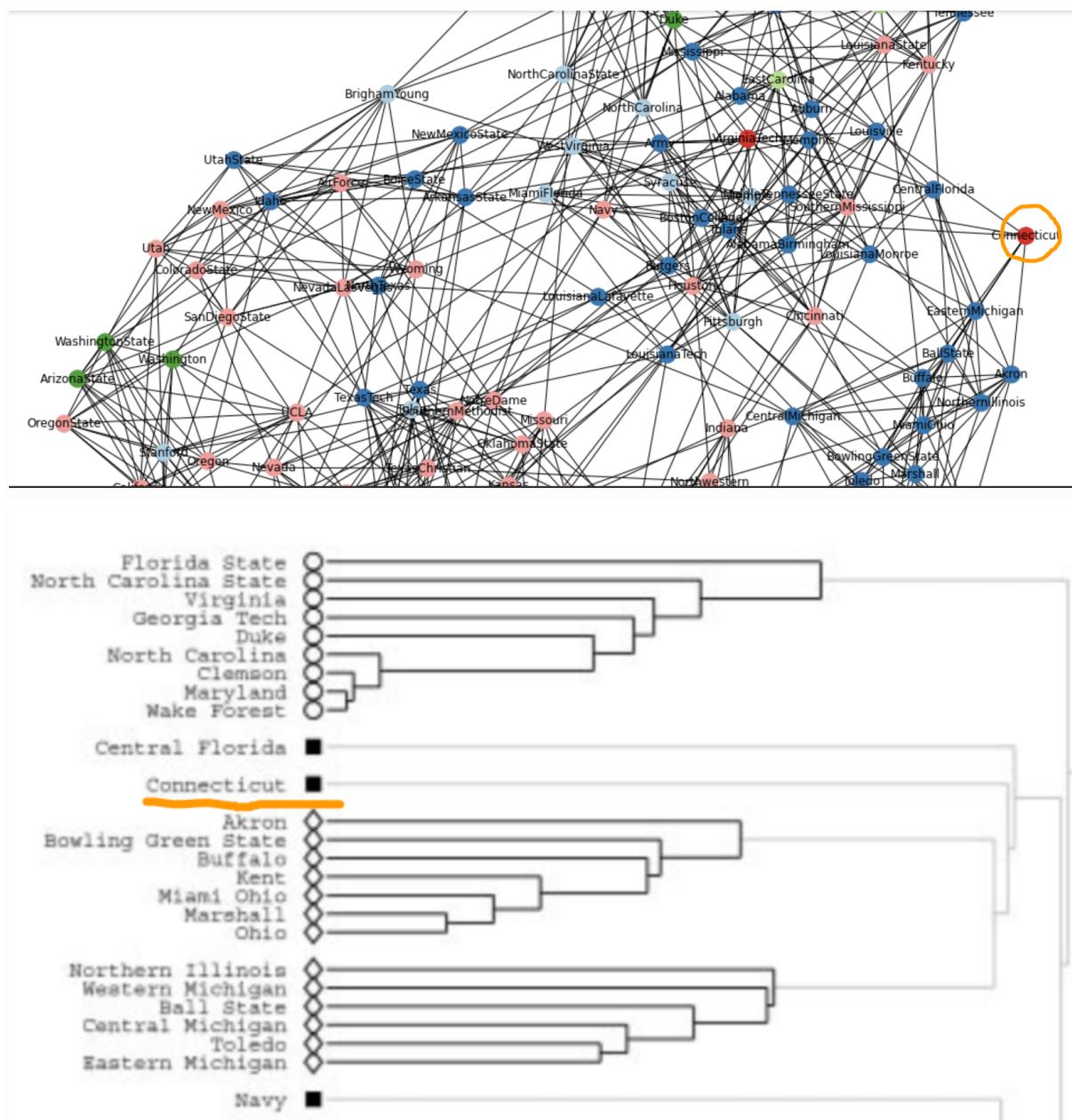
### 3. 그래프 시각화

거의 모든 팀이 컨퍼런스에서 같은 컨퍼런스에 속한 다른 팀과 올바르게 그룹화되어 있음을 그래프를 통해서 알 수 있다.

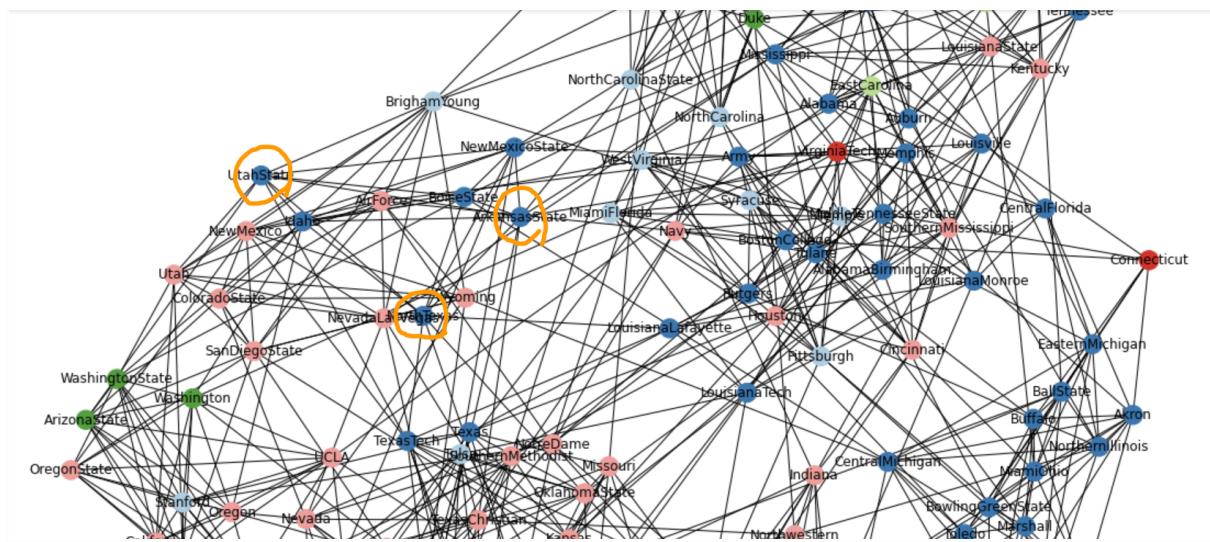


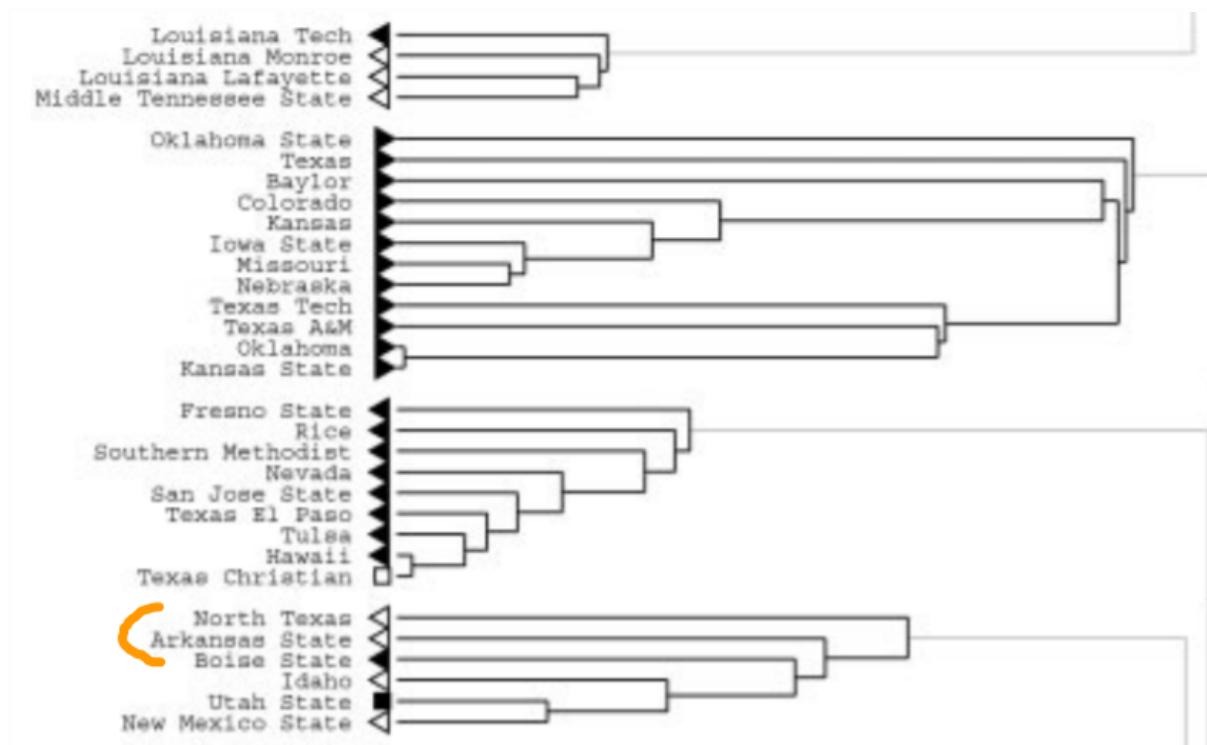
#### 4. 그래프 분석 및 설명

그러나 아래 그림처럼 어떤 회의에도 속하지 않는 독립 팀이 몇 개 있는데, 이들은 가장 밀접하게 연관된 회의와 함께 분류되는 경향이 있다.



또한 같은 컨퍼런스에 속하였지만 다른 네트워크(색)으로 구분이 되는 팀들이 있는데 이는 지리적으로는 비슷하나 다른 컨퍼런스에 있는 팀들과 게임을 더 많이 했기 때문이다. 예를 들어, 선벨트 컨퍼런스는 두 부분으로 나뉘고 서부 체육 컨퍼런스의 멤버들과 함께 그룹화되는데, 이것은 선벨트 팀들이 서부 컨퍼런스 팀과 거의 많은 게임을 했기 때문에 일어난다.





## 5. Reference

보고서에서 작성한 미국 대학 미식축구 (Schedule of Division I games for the 2000 season)의 대진표는 Proceedings of the National Academy of Sciences of the United States of America에서 참고함.