

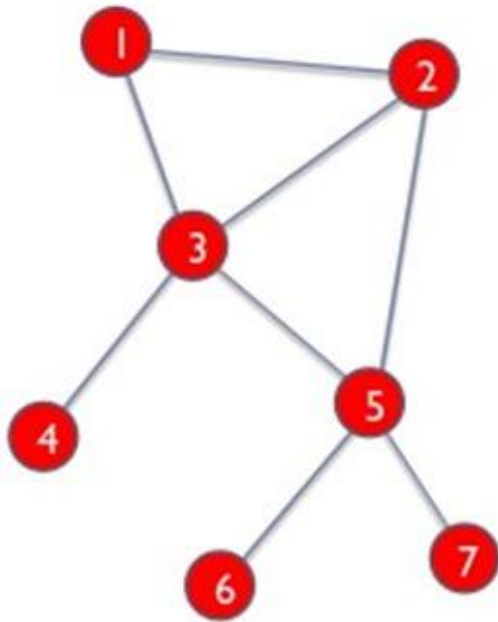


Lecture 8

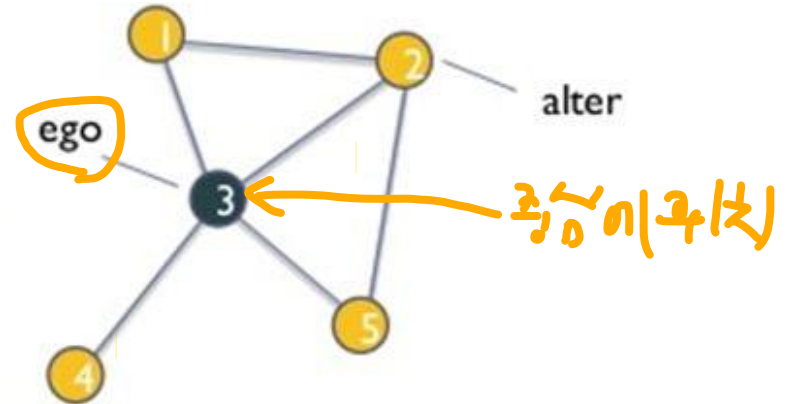
사회연결망분석 (2)

에고 네트워크

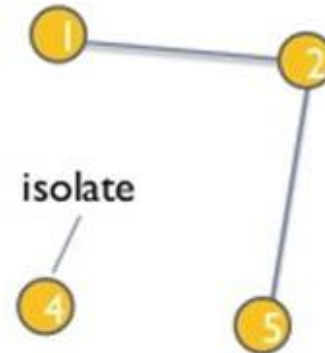
'whole' network*



3's ego network



3's ego network without ego**



사건 매트릭스

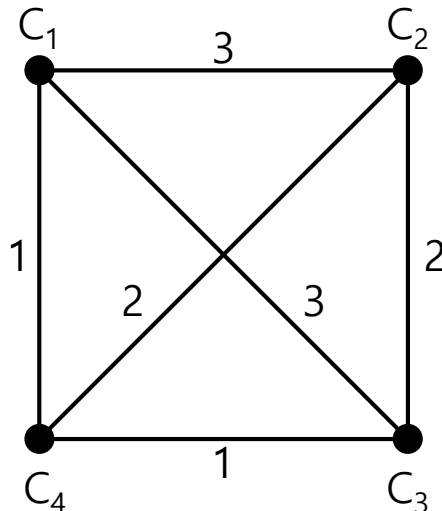
	A	B	C	D	E
C_1	1	1	1	0	1
C_2	1	1	1	0	1
C_3	0	1	1	1	0
C_4	0	0	1	0	1



인접도 매트릭스

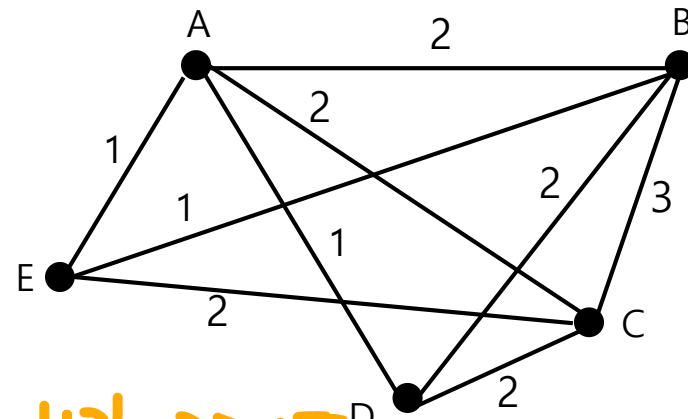
기업 vs. 기업

	C ₁	C ₂	C ₃	C ₄
C ₁	-	3	3	1
C ₂	3	-	2	2
C ₃	3	2	-	1
C ₄	1	2	1	-



사외이사 vs. 사외이사

	A	B	C	D	E
A	-	2	2	1	1
B	2	-	3	2	1
C	2	3	-	2	2
D	1	2	2	-	0
E	1	1	2	0	-



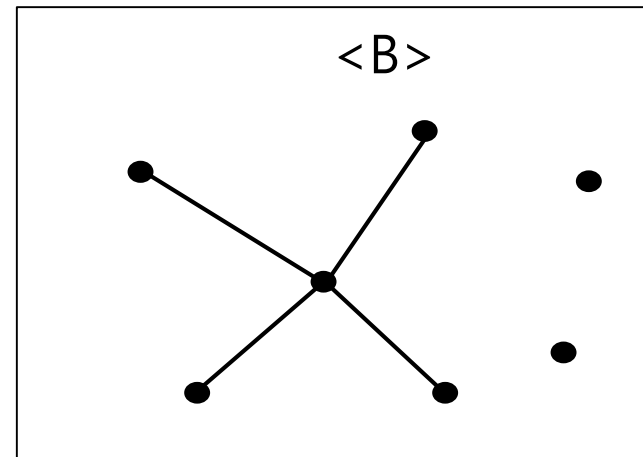
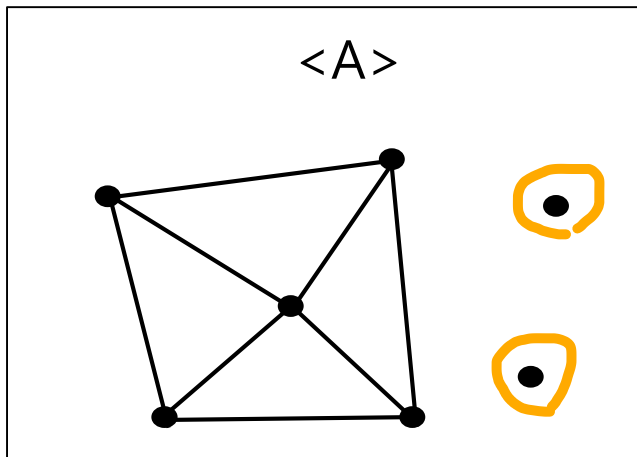
계량 그래프



밀도 측정

❖ 절대적 포괄성 = 전체 점의 수 - 고립된 점의 수 5

❖ 상대적 포괄성 = $\frac{\text{연결된 점의 수}}{\text{네트워크 내 전체 점들의 수}}$ $\frac{5}{7}$



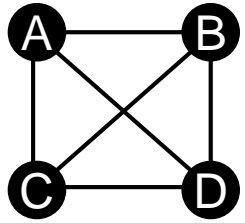
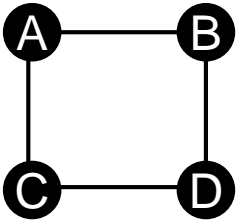
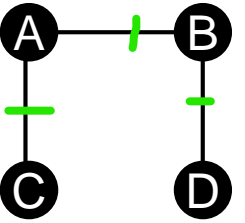
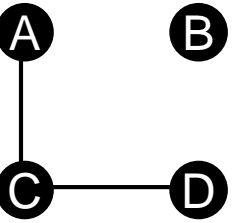
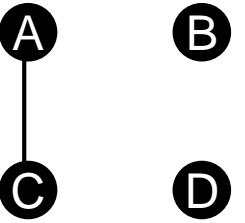
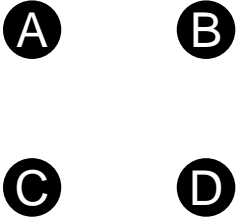
- 상대적 포괄성: 5/7
- 한 그래프의 연결정도의 총합 = $(2 * \text{총 라인의 수})$
- A의 연결정도 총합: 16 ($=2*8$)
- B의 연결정도 총합: 8 ($=2*4$)



- ❖ **밀도**: 네트워크 내 실제 연결선의 수가 그 점들이 가질 수 있는 **최대한의 연결에서 차지하는 비중**

$$\text{밀도(density)} = \frac{L}{g(g-1)/2}$$

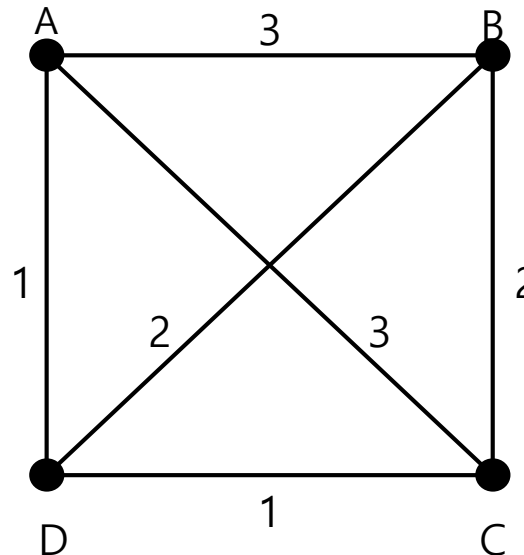
* L: 네트워크 내 연결선의 수, g: 점의 총수

구 분						
연결된 점의 수	4	4	4	3	2	0
상대적 포괄성	1.0	1.0	1.0	0.75 $\frac{3}{4}$	0.5 $\frac{2}{4}$	0
연결정도의 합	12	8	6 3×2	4	2	0
연결선의 수	6	4	3	2	1	0
밀도	1.0	0.667(4/6)	3/6	2/6	1/6	0



방향, 계량 네트워크에서의 밀도 계산

- ❖ 방향네트워크 밀도 = $\frac{L}{g(g-1)}$ "나누기 2 x"
- ❖ 계량네트워크 밀도 = $\sum_{l=1}^n v_l / g(g-1)$
 - v_l 는 라인 l 의 값, g 는 네트워크에 존재하는 노드의 수



합 = 12,,

가정: 한 네트워크에서 어떤 두 점 간의 최대 가능한 다중성은 3



계량값을 가진 그래프

- ❖ 최대 발생할 수 있는 총 다중성: $3 \times 6 = \underline{\underline{18}}$
- ❖ 밀도: $0.667 (= 12/18)$
 - 각 라인에 부여된 계량값의 합이 12이므로
- ❖ 만약, 이 그래프에서 **관계의 계량값을 고려하지 않는다면** 밀도의 값은
$$\frac{6}{4(4-1)/2} = 1.0 \,, \Rightarrow \text{완전연결}$$
- ❖ 만약, **최대 가능 다중성의 값이 5라고 하면**, 밀도의 값은 $0.4(12/30)$ 가 되는데, 이는 위의 0.667보다 낮아진 값임 5×6



❖ 중심성 분석(centrality analysis)

❖ 중심의 개념

- 가시성이 높을 경우 (주목 받는 경우, 매력적인 경우)
- 다른 사람과 관계를 가지는 정도가 매우 높은 경우 (마당발, 교류가 많은 경우)
- 네트워크 내 구조적 위치가 중개자의 위치인 경우 (다른 사람들간 교류의 bridge가 되기 때문)



여러 가지 중심성의 개념들

- ❖ 내향중심성(in-centrality)
- ❖ 외향중심성(out-centrality)
- ❖ 포인트 중심성(point centrality)
- ❖ 그래프 중심성(graph centrality)
- ❖ 로컬 중심성
- ❖ 글로벌 중심성



내향 vs. 외향중심성

❖ 내향중심성 (in-centrality)

- 교류방향이 외부에서 자신 쪽으로 들어오는 경우
- '명예(prestige)'

❖ 외향중심성 (out-centrality)

- 교류방향이 자신에서 외부 쪽으로 나가는 경우
- '마당발'



포인트 vs. 그래프 중심성

❖ 포인트 중심성 (point centrality)

- 특정한 점(노드)이 네트워크 내의 중심에 위치하는 정도를 표현하는 지표
- 한 점의 중요도 혹은 탁월함을 의미하는 개념

❖ 그래프 중심성 (graph centrality)

- 전체 그래프의 중심화 경향을 나타내는 개념
- 네트워크 전체가 가진 종합적인 중심에의 응집을 의미
- 집중도(centralization)



로컬 vs. 글로벌 중심성

❖ 로컬 중심성 크기 고려 X

- 한 점을 중심으로 주변의 점들과 얼마나 많이 연결되어 있는가에 의해 파악됨
- 한 점과 직접적으로 연결되어 있는 점의 수만을 가지고 측정함
- 네트워크에서 자신과 연결된 로컬 영역에 한정하여 중심성 측정

❖ 글로벌 중심성 크기 고려 O

- 특정 한 점이 전체 네트워크의 중심에 위치하는 정도를 나타내는 개념
- 한 점과 네트워크의 전체 점들간의 거리에 의해 측정됨



연결정도, 근접, 매개중심성

❖ 연결정도(degree)중심성 인스타 팔로워(?)

- 다른 점(노드)과의 연결된 정도를 중심으로 보는 개념
- 한 점에 연결된 다른 점의 수로 측정

❖ 근접(closeness)중심성

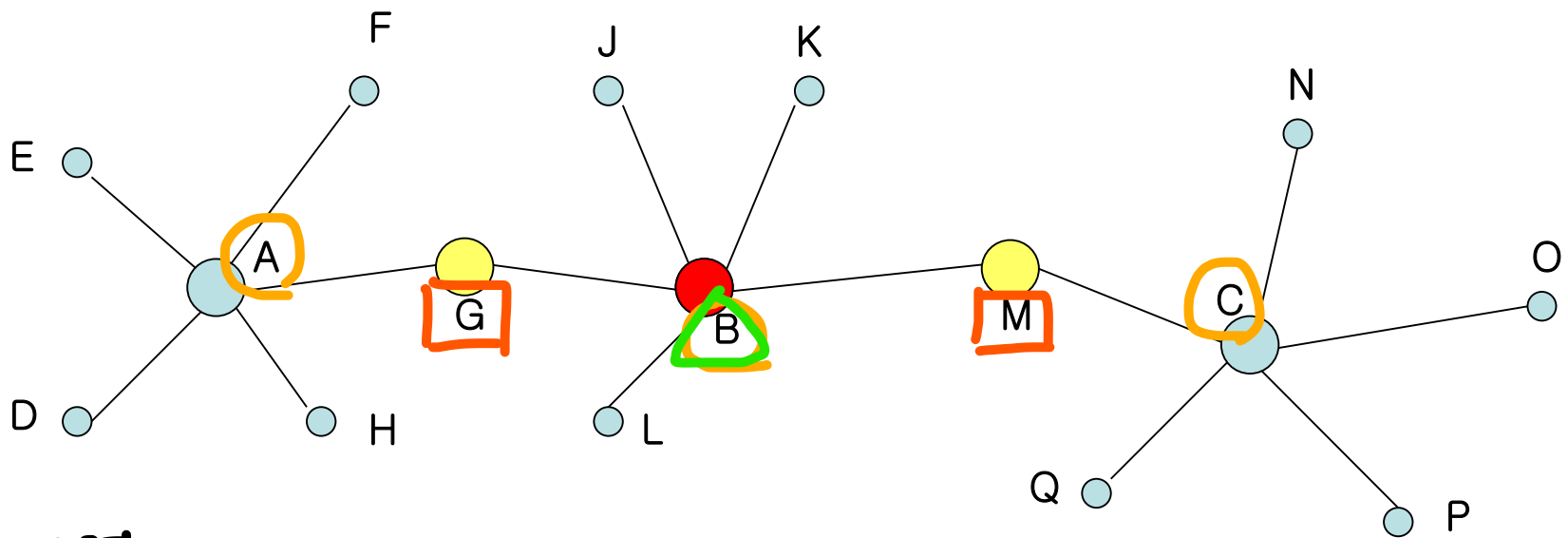
- 한 점이 다른 점에 얼마 만큼 가깝게 있는가를 말하는 개념
- 두 점 사이의 거리(distance)가 핵심 개념

❖ 매개(betweenness)중심성

- 한 점이 다른 점들과의 네트워크를 구축하는데 얼마나 중개자 혹은 다리(bridge) 역할을 수행하느냐를 측정하는 개념
- 중개 역할에 초점을 맞추어 '중심'으로 간주하는 특징을 가짐



중심성과 네트워크



○ 연결

△ 근접

□ 매개



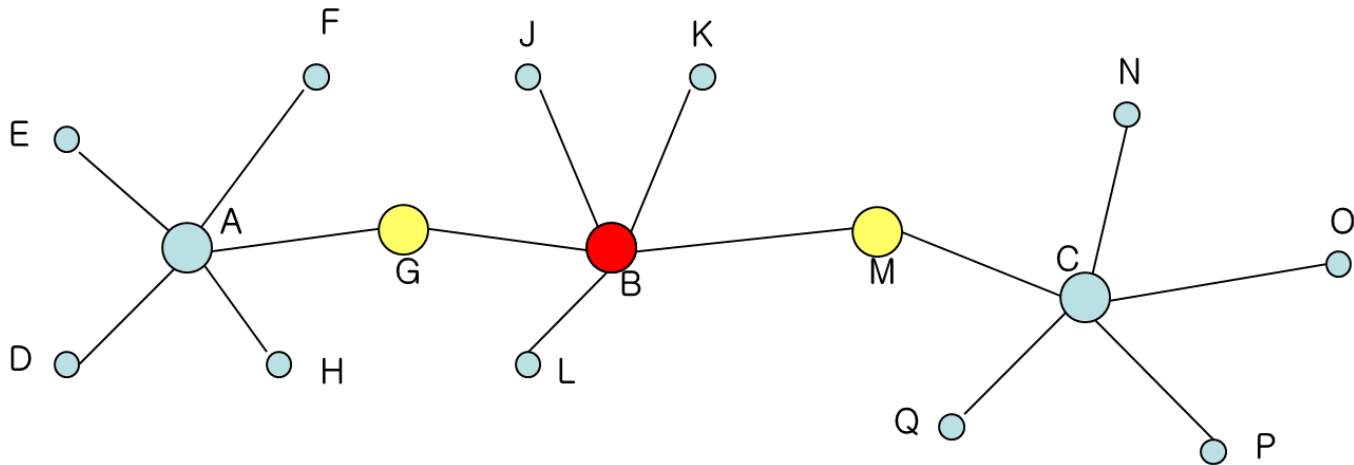
연결정도중심성(degree centrality)

- ❖ 한 점이 다른 점들과 얼마만큼 직접 연결되었는지에 의해 중심 정도를 평가하는 방법
 - '친구의 친구'는 고려하지 못하므로 로컬중심성의 의미가 강함
- ❖ 절대적 중심성
 - 한 점 주변에 직접적으로 연결되어 있는 이웃점의 '절대적' 수로 측정됨
 - 네트워크 규모의 차이를 반영하지 못함 x
 - [그림] A, C: 5, G, M: 2
- ❖ 상대적 중심성
 - 어떤 한 점 A의 상대적 중심성은 그 A에 연결된 점의 수를 전체 가능한 점의 수로 나눈 값
 - $C_A = A\text{의 절대적인 연결정도} / (\text{네트워크 내 전체 점의 수} - 1)$



연결정도중심성

구분		A,C	B	G,M	J,K,L	다른 점들
로컬중심성	절대적	5	5	2	1	1
	상대적	5/15	5/15	2/15	1/15	1/15
글로벌중심성		43	33			



근접중심성(closeness centrality)

❖ 근접중심성의 계산

- 각 점간의 거리를 근거로 하여 중심성을 측정하는 방법
- 네트워크 내 직간접적으로 연결된 모든 점들 간의 거리를 계산하여 중심성을 측정함
- **간접적 연결**: 중간에서 매개하는 점(노드)을 거쳐서 연결되는 경우
- 근접중심성 지수가 높을수록 다른 점들과 가까이 위치하고 있다는 의미
- 한 점의 글로벌 중심성 측정 가능
- 다른 점들이 그 점과 연결될 수 있는 거리들을 모두 더한 값의 역수

$$C_i = \left[\sum_{j=1}^n d_{ij} \right]^{-1}$$

“절대적 근접중심성”

d_{ij} : 두 점 i, j 를 잇는 가장 짧은 경로

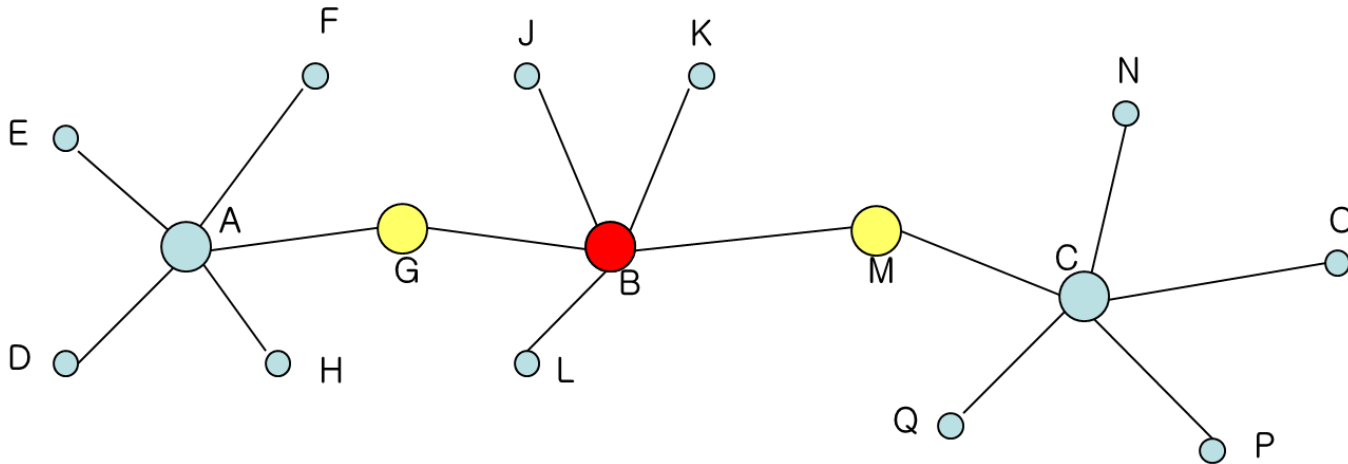


근접중심성

❖ 상대적 근접중심성

$$C_i = (g - 1) \left[\sum_{j=1}^n d_{ij} \right]^{-1}$$

37 $\begin{bmatrix} 2 & 1 \times 2 \\ 16 & 2 \times 8 \\ 3 & 3 \times 1 \\ 16 & 4 \times 4 \end{bmatrix}$



구분		A,C	B	G,M	J,K,L	다른 점들
근접중심성	절대적	0.0232(=1/43)	0.0303(=1/33)	0.027(1/37)	0.0212(=1/47)	0.0175(=1/57)
	상대적	0.348 (=(16-1)*0.0232)	0.454 (=(16-1)*1/33)	0.405	0.319 (=(16-1)*0.0212)	0.263 (=(16-1)*0.0175)



매개중심성(betweenness centrality)

❖ 매개중심성의 계산

- 네트워크 내에서 한 점이 담당하는 매개자 혹은 중재자 역할의 정도로서 중심성을 측정하는 방법
- 높은 매개중심성을 가질 때 Broker, Gatekeeper 역할
- 중재역할의 중요성이 클수록 의사 소통을 제어 할 수 있는 통제력이 커지고, 다른 행위자들의 이에 대한 의존성도 커지게 됨
- 점 X와 Z 사이에 존재하는 Y의 매개중심성 = 네트워크 상에서 X와 Z를 연결하는 가장 짧은 경로 중에서 Y가 포함되어 있는 경로의 비율로 측정함
- X와 Z 사이에 존재하는 짧은 거리경로 중 Y를 통과해야 하는 경로숫자의 비율

$$C_B(i) = \sum_{j < k} g_{jk}(i) / g_{jk}$$

g_{jk} : 특정 두 점(j와 k) 사이에 존재하는 최단거리경로의 경우의 숫자

$g_{jk}(i)$: 두 점 j와 k(j≠k) 사이에 존재하는 점 i를 경유하는 횟수

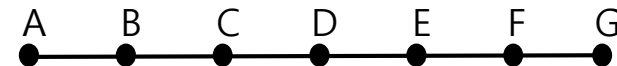
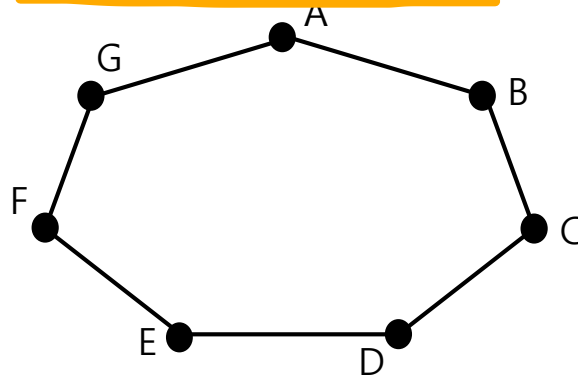
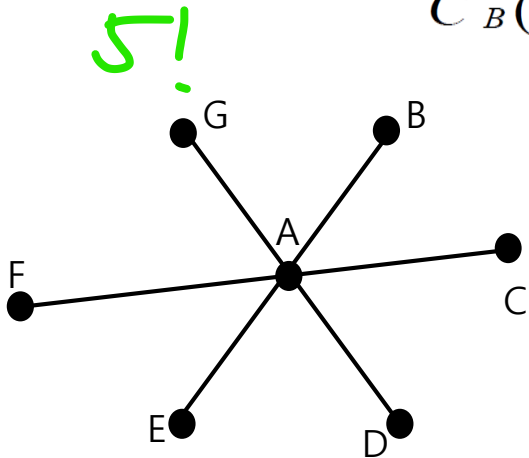


매개중심성

❖ 상대적 매개중심성

$$C'_B(i) = \frac{\sum_{j < k} g_{jk}(i) / g_{jk}}{[(g-1)(g-2)/2]}$$

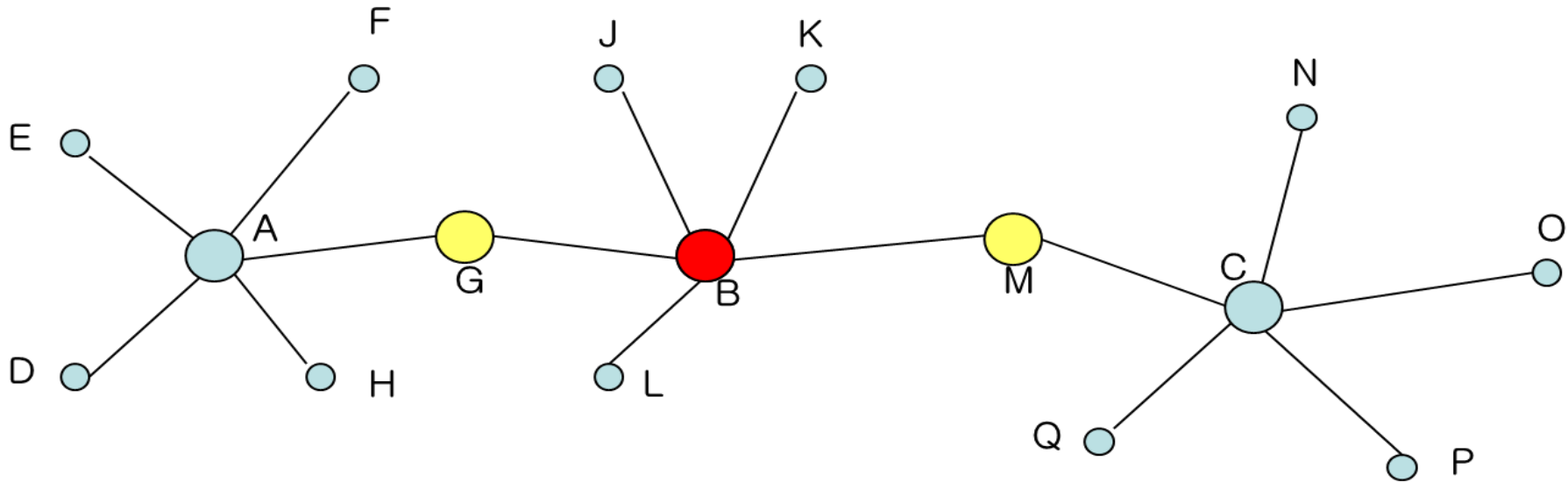
$[(g-1)(g-2)/2]$: 최대 가능한 매개중심성 값
 $(= {}_{g-1}C_2 = (g-1)! / [(g-1-2)! \times 2!])$



구 분		A	B	C	D	E	F	G
스타형	절대적	15	0	0	0	0	0	0
	상대적	1.0	0	0	0	0	0	0
원 형	절대적	3	3	3	3	3	3	3
	상대적	0.2	0.2	0.2	0.2	0.2	0.2	0.2
라인형	절대적	0	5	8	9	8	5	0
	상대적	0	0.33	0.533	0.6	0.533	0.33	0



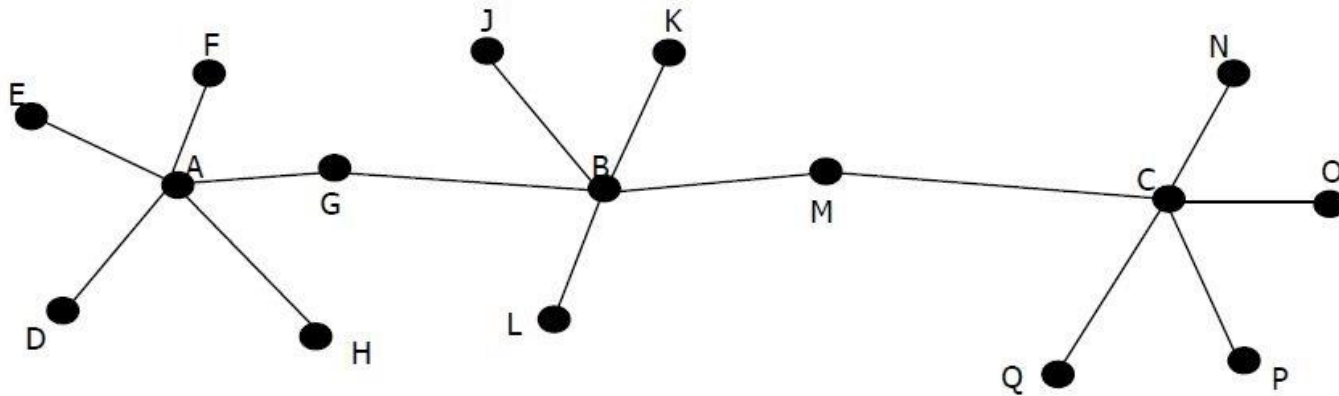
매개중심성



구분		A,C	B	G,M	J,K,L	다른 점들
매개중심성	절대적	50	75	50	0	0
	상대적	$0.4762 (50/(16-1)(16-2)/2)$	$0.7143 (75/(16-1)(16-2)/2)$	$0.4762 (50/(16-1)(16-2)/2)$	0	0



연결정도, 근접, 매개중심성 비교

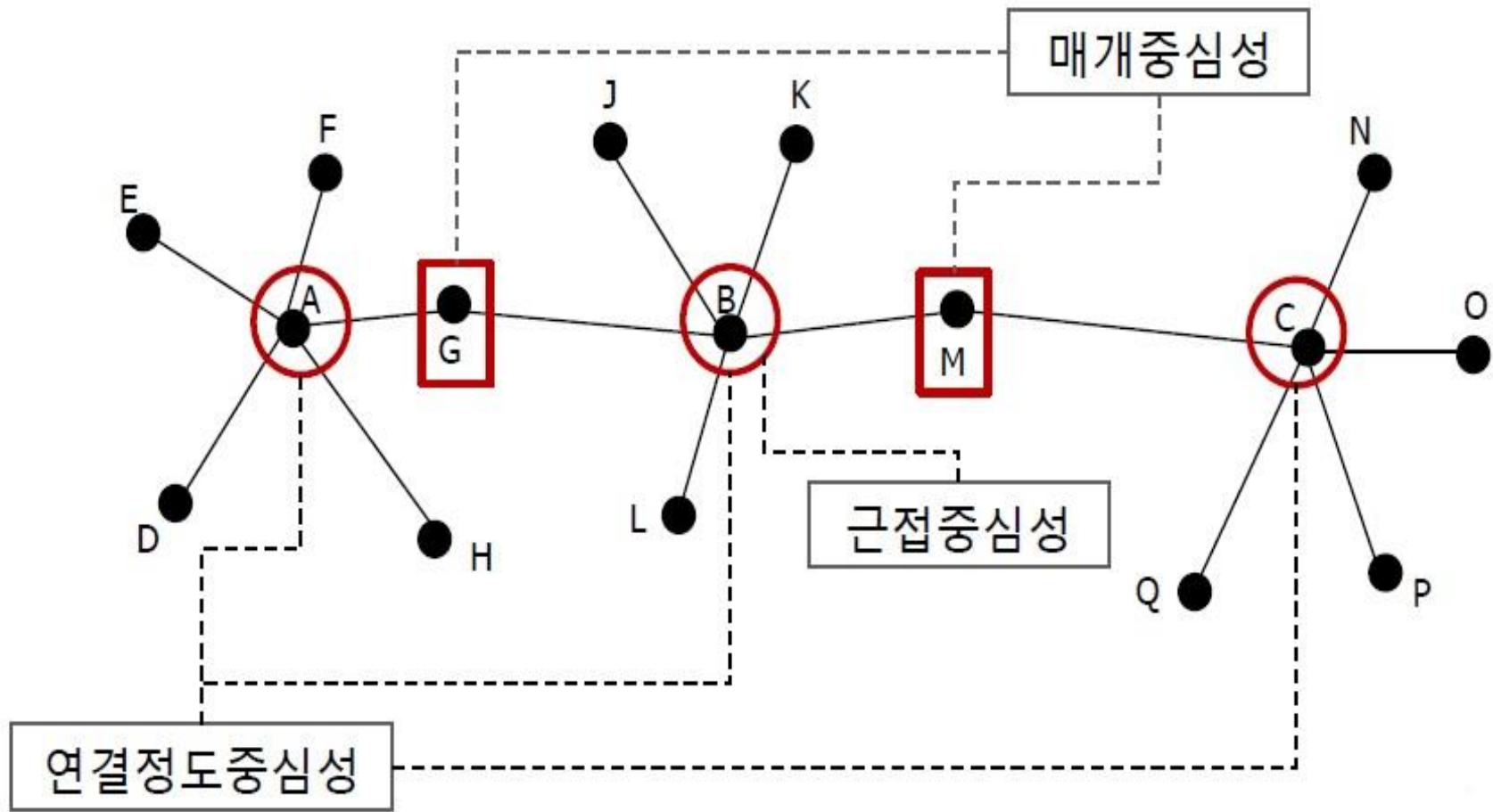


	구분		A,C	B	G,M	J,K,L	다른 점들
연결정도 중심성	로컬중심성	절대적	5	5	2	1	1
		상대적	0.33	0.33	0.13	0.07	0.07
	글로벌중심성 ✕		43	33	37	47	57
근접 중심성	절대적		0.0232	0.0303	0.0270	0.0212	0.0174
	상대적		0.348	0.454	0.405	0.319	0.263
매개 중심성	절대적		50	75	50	0	0
	상대적		0.4762	0.7143	0.4762	0	0



연결정도, 근접, 매개중심성 비교

❖ 중심성과 네트워크



연결정도, 근접, 매개중심성 비교

❖ 연결정도중심성: **인기**

- 한 네트워크에서 한 개인이 담당하는 정보교류의 정도를 표현

❖ 근접중심성, 매개중심성:

- 한 개인의 네트워크 내에서의 정보교류에 대한 통제능력을 표현
- 정보통제 측면에서 매개중심성이 근접중심성보다 현저한 통제력을 보이는 지표임
 - 매개중심성은 다른 행위자간 매개에 의존하는 측정방법이기 때문
- 두 중심성은 소수의 특정 행위자들이 정보통제를 담당하는 정도를 포착하는 특징을 가짐

❖ 한편, **집중도**는 한 네트워크에서 총체적으로 행위자들간 정보흐름의 활성화 정도를 표현함



- ❖ 한 네트워크 전체가 한 가지 '중심'으로 집중되는 정도를 표현하는 지표
 - 가장 중심적인 점의 중심성 값과 다른 점들의 중심성 값 간의 차이를 보는 것
 - 네트워크 상의 실제 이 차이의 합과 이론적으로 가능한 최대 차이의 합 간의 비율로 계산
 - 가장 중심적인 점과 다른 모든 점들의 중심성 점수들간의 차이를 각각 구하여 이를 모두 합한 다음, 이것을 논리적으로 가능한 최대값으로 나누는 것
 - 집중도를 측정하는 방법은 중심성을 측정하는 세 가지 방법에 각각 대응해서 존재함
- ❖ 중심성 vs 집중도
 - 중심성은 한 점의 입장을, 집중도는 전체 네트워크의 입장을 가짐
- ❖ 밀도 vs 집중도
 - 밀도와 집중도는 반비례 추세를 가짐
 - 밀도는 구성원간 연결관계의 수로 추론되는 평균(mean) 개념
 - 집중도는 구성원간 연결관계의 분산(variance) 개념



연결정도 집중도 (Degree Centralization)

- ❖ 전체 네트워크의 집중화 정도를 측정하기 위해, 각 점들간 연결정도에 의존하는 측정방법

$$C_D = \frac{\sum_{i=1}^g [C_D(n^*) - C_D(n_i)]}{\text{Max} \sum_{i=1}^g [C_D(n^*) - C_D(n_i)]} = \frac{\sum_{i=1}^g [C_D(n^*) - C_D(n_i)]}{[(g-1)(g-2)]}$$

$C_D(n^*)$: 네트워크 내 가장 높은 연결정도중심성 값

$C_D(n_i)$: 어떤 한 점 i 의 연결정도중심성의 값

$[(g-1)(g-2)]$: 한 네트워크에서 이론적으로 가능한 연결정도중심성의 값 차이의 최대값

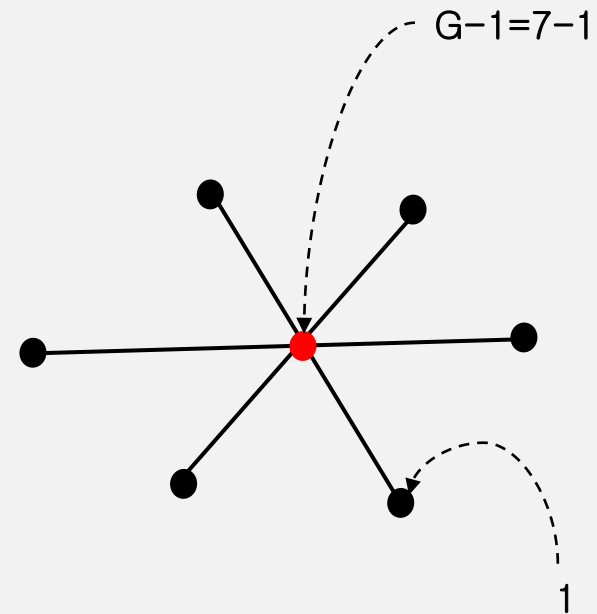
- ❖ 분자: 네트워크 내의 가장 큰 연결정도 중심성 값과 나머지 $(g-1)$ 개 다른 점들의 연결정도 중심성 간의 차이를 합산한 것
- ❖ 분모: 이론적으로 가능한 이러한 차이합의 최대값 (스타형에서의 값)



연결정도 집중도 (Degree Centralization)

❖ 스타 네트워크의 사례

- 연결정도 중심성: 중앙 노드 = $(g-1)$, 나머지 노드 = 1
- 가장 큰 연결정도 중심성 - 나머지 노드의 연결정도 중심성 = $(g-1)-1 = g-2$
- 스타 네트워크에서 이 차이는 $(g-1)$ 번 발생하므로 최대 가능한 연결정도 중심성의 차이의 합 = $(g-2)(g-1) \rightarrow$ 분모



- 최대 중심성과 개별 중심성의 차이 = $(g-1)-1=g-2=7-2=5$
- 최대 중심성과 개별 중심성의 차이의 합 = $(g-2) \times (g-1) = (7-2) \times (7-1) = 30$
- 그룹 연결정도 집중도 = $30/30 = 1$



근접 집중도(Closeness Centralization)

- ❖ 각 점의 거리에 기반을 둔 측정방법 네트워크 전체의 근접 집중도는 근접 중심성(closeness centrality)을 사용하여 계산

$$C_c = \frac{\sum_{i=1}^g [C_c(n^*) - C_c(n_i)]}{[(g-2)(g-1)/(2g-3)]}$$

$C_c(n^*)$: 네트워크 내 근접중심성에서 가장 큰 값

$C_c(n_i)$: 어떤 한 점 i 의 근접중심성 지수

$[(g-1)(g-2)/(2g-3)]$: 한 네트워크에서 이론적으로 가능한 근접중심성 값 차이의 최대값

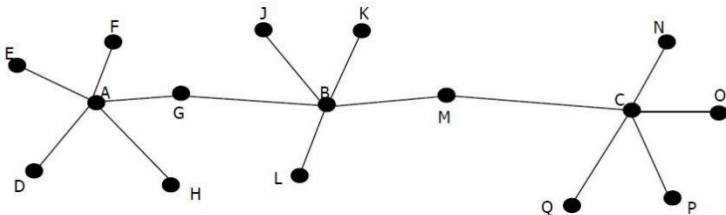


근접 집중도(Closeness Centralization)

❖ 스타 네트워크의 사례

- 중앙 노드: 최단경로거리의 합 = $(g-1)$, 상대적 근접중심성 = $(g-1) \times 1/(g-1) = 1$
- 나머지 노드들: 중앙 노드와 경로거리 1, 다른 노드들과의 경로거리 2 이므로 최단경로거리의 합 = $1 + 2(g-2) = (2g-3)$, 상대적 근접중심성 = $(g-1) \times 1/(2g-3) = (g-1)/(2g-3)$
- 가장 큰 상대적 근접중심성과 나머지 노드들의 상대적 근접중심성 차이 = $1 - (g-1)/(2g-3)$
- 그 차이의 합 = $(g-1) \times (1 - (g-1)/(2g-3)) = ((g-2)(g-1))/(2g-3)$

❖ 아래 그림의 근접 집중도 = $[\{(0.454 - 0.348) \times 2\} + \{(0.454 - 0.405) \times 2\} + \{(0.454 - 0.319) \times 3\} + \{0.454 - 0.263\} \times 8] / \{16 - 2\}(16 - 1) / (2 \times 16 - 3) = 0.312$



근접 중심성	구분	A,C	B	G,M	J,K,L	다른 점들
	절대적	0.0232	0.0303	0.0270	0.0212	0.0174
	상대적	0.348	0.454	0.405	0.319	0.263



매개 집중도(Betweenness Centralization)

- ❖ 각 점들의 매개성을 기초로 하여 한 네트워크의 집중도 정도를 파악하는 개념

$$C_B = \frac{\sum_{i=1}^g [C_B(n^*) - C_B(n_i)]}{[(g-1)^2(g-2)]/2}$$

?

$C_B(n^*)$: 한 네트워크에서 얻는 매개중심성의 최대값

$C_B(n_i)$: 어떤 한 점 i 의 매개중심성 지수

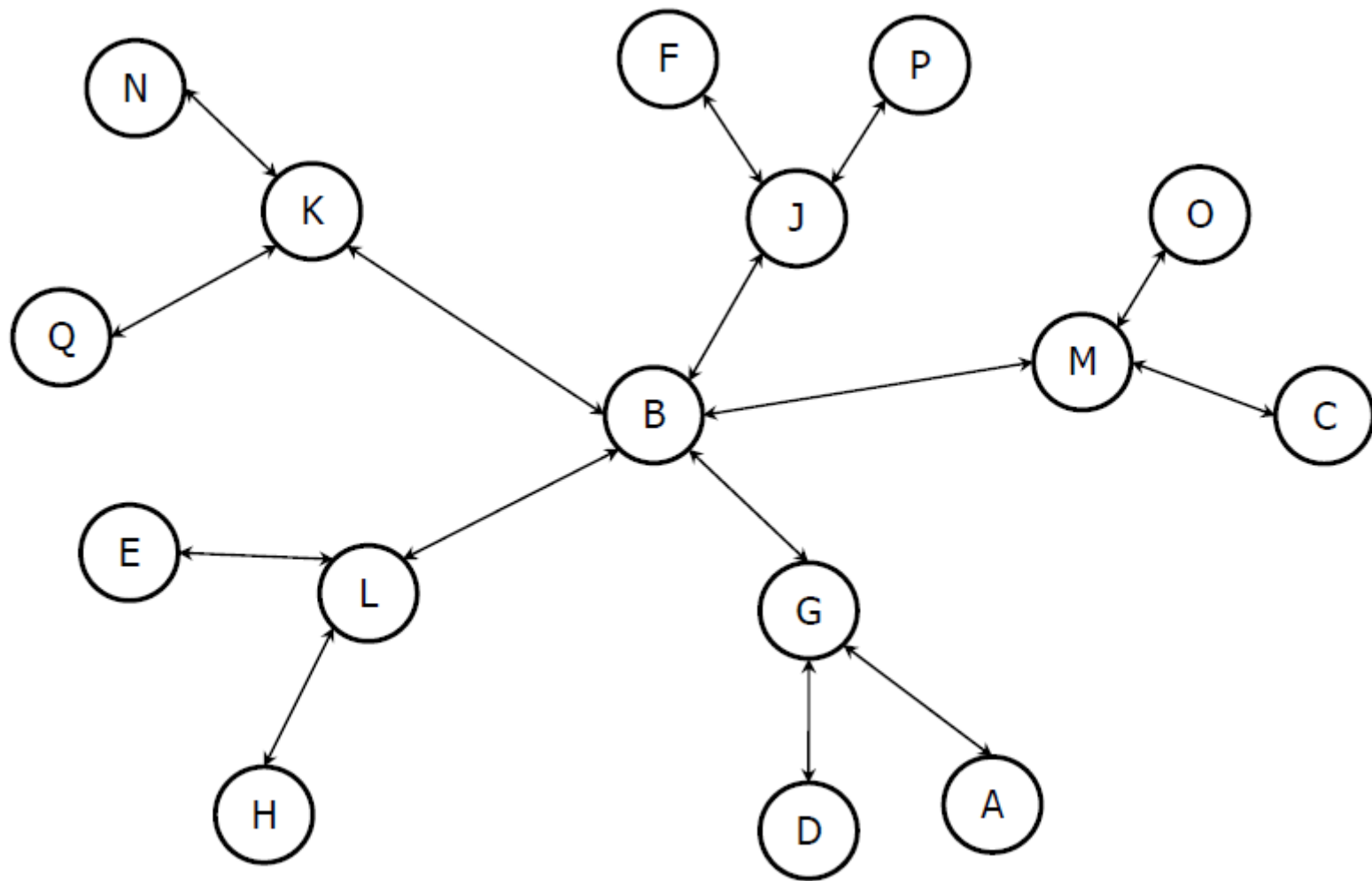
합의

$[(g-1)^2(g-2)]/2$: 네트워크 내에서 이론적으로 가능한 매개중심성 값 차이의 최대값 (최대 매개중심성 값 $= (g-1)(g-2)/2$ 이므로 분모의 매개중심성 차이의 합은 $\{(g-1)(g-2)/2 - 0\} \times (g-1)$)

- ❖ 0~1 (스타형그래프→1, 원형그래프→0) "스타형과 내가 얼마나 하야 나눈지"



집중도 계산 예



연결정도집중도 계산 예

- ❖ 가장 큰 연결정도중심성: B (5)
- ❖ 가장 변방에 있는 점들의 연결정도중심성: A,D,C,O,F,P,N,Q,E,H (1)
- ❖ 중간에 있는 점들의 연결정도중심성: L,G,M,J,K (3)
- ❖ 분자 계산
 - 최대값인 5로부터 모든 점들의 중심성을 뺀 차이를 더한 값
$$= (5-1) \times 1 + (5-3) \times 5 = 50$$
- ❖ 분모 계산
 - 논리적으로 가장 큰 차이를 합한 값
 - 16개로 이루어진 연결망에서 한 점이 논리적으로 가질 수 있는 최대값은 스타형 연결망에서의 15
 - 스타형 연결망에서 변방에 놓은 15개 점들의 중심성은 1
 - 차이의 합: $(15-1) \times 15 = 210$
- ❖ 따라서 연결정도집중도는 $50/210 = 0.238$

