



## 저작자표시 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.
- 이차적 저작물을 작성할 수 있습니다.
- 이 저작물을 영리 목적으로 이용할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#) 

공학박사학위논문

감마분포로 추출한 허브 네트워크와  
확산 및 구조 신뢰성 해석

Spreading and Structural Reliability Analysis  
by Extraction of Hub Network  
with Gamma Distribution

2014년 2월

서울대학교 대학원

생태조경·지역시스템공학부 지역시스템공학전공

이 형 진



감마분포로 추출한 허브 네트워크와  
확산 및 구조 신뢰성 해석

Spreading and Structural Reliability Analysis  
by Extraction of Hub Network  
with Gamma Distribution

지도교수 이 정 재

이 논문을 공학박사 학위논문으로 제출함

2013년 12월

서울대학교 대학원

생태조경·지역시스템공학부

지역시스템공학 전공

이 형 진

이형진의 공학박사 학위논문을 인준함

2014년 1월

위 원 장 \_\_\_\_\_ (인)

부위원장 \_\_\_\_\_ (인)

위 원 \_\_\_\_\_ (인)

위 원 \_\_\_\_\_ (인)

위 원 \_\_\_\_\_ (인)



## 국문초록

실세계에서 계(System)는 전체 구조를 대표해주는 허브로 구성된 계와 배경처럼 모든 곳에서 동일한 무작위 계가 혼합되어 있다. 만약 대상에서 전체 구조를 대표할 수 있는 허브 계를 추출할 수 있다면 해석의 정확성과 효율성을 동시에 향상시킬 수 있을 것이다.

연속체로 구성된 계를 간단한 이산화 모델로 변화시켜 해석하는 경우의 네트워크도 앞에서 논한 바와 같이 절점이 불균등한 분포를 가지는 허브 네트워크와 절점이 균등하게 분포된 무작위 네트워크로 구분된다. 두 네트워크는 서로 다른 특성을 가지며 해석방법도 다르다. 따라서 본 연구에서는 실제 문제에서 대표성 있는 허브 네트워크를 추출하고, 그 허브 네트워크를 거시적 관점의 취약성 분석, 미시적 관점의 군집화 분석 등의 해석을 확산 및 구조 신뢰성 문제에 적용하여 고찰하였다.

모든 네트워크는 절점을 연결하는 연결선의 개수를 변수로 하는 경우 고유한 연결선의 분포를 가지게 된다. 이를 감마분포로 나타내면 분포의 고유값에 해당하는 형상계수와 규모계수를 구할 수 있다. 그리고 이 두 개의 고유값을 이용하면 대상 네트워크에서 허브 네트워크를 추출할 수 있는 임계치를 산정할 수 있다.

추출된 허브 네트워크는 불균등한 분포를 가지기 때문에 연결성에 따라서 절점들의 중요도가 달라진다. 많은 연결선을 가지는 허브 절점은 전체 네트워크의 규모에 결정적인 영향을 미치므로, 네트워크의 생성, 소멸과 같은 외적 변화를 살피는데 유용하다. 허브 탐색 방법을 가축전염병의 확산과 소멸에 적용하여 효율적으로 확산을 차단할 수 있는 방안에 대해서 고찰하였다.

외적으로 고정된 허브 네트워크에서 연결선의 속성이 확률적으로 변동하는 경우, 컨볼루션을 이용한 군집화를 통해서 전체 네트워크의 내부 변화를 고찰할 수 있다. 뼈대 구조에서 파괴에 영향을 주는 허브 네트워크를 군집화하여 효율적으로 구조 신뢰성을 적용하는 방안을 고찰하였다.

허브 네트워크의 추출과 취약성 분석 및 군집화 등과 같은 해석 방법을 구조 문제와 확산 문제에 각각 적용한 결과, 연속체에서 최적 구조 추출을 위한 수행지수와 비교할 때 각 단계별 변형에너지와 부피를 산정할 필요 없이 연결선 수 분포의 분석만으로 약 98% 이상의 정확도를 나타냈다. 그리고 뼈대 구조의 구조신뢰성 해석시 몬테카를로 해석과 같이 1,000,000번의 연산 없이 각 외력에 대한 한 번의 군집화 분석만으로 Beam, Sway 모드 파괴확률이 95%이상의 정확도를 가지는 것을 확인하였다. 가축전염병의 확산 문제에서는 허브네트워크를 추출하고 취약성 분석을 수행하여 전체의 12% 농가만을 차단하여 기존의 87% 농가를 차단하는 대책과 동일한 효과를 낼 수 있음을 확인하였다.

주요어 : 허브 네트워크, 감마분포, 취약성, 군집화,  
컨볼루션, 뼈대구조

학 번 : 2008-30922

# 목 차

제 1 장 서 론 .....	1
1.1 연구배경 .....	1
1.2 연구사 .....	3
1.2.1 복잡계 네트워크 .....	3
1.2.2 네트워크 추출 .....	5
1.2.3 중심성 분석 .....	6
1.2.3 군집화 분석 .....	7
1.3 연구목적 및 범위 .....	8
1.2.1 감마분포를 이용한 허브 네트워크 추출 .....	8
1.2.2 허브 네트워크의 취약성 분석 .....	9
1.2.3 확률을 이용한 허브 네트워크 .....	10
1.4 연구구성 .....	10
참고문헌 .....	13
제 2 장 감마분포를 이용한 허브 네트워크 추출 .....	17
2.1 서 론 .....	17
2.2 혼합계의 네트워크 표현 .....	19
2.3 감마 분포의 형상계수 시뮬레이션 .....	22
2.3.1 무작위 네트워크 .....	25
2.3.2 허브 네트워크 .....	26
2.3.3 혼합 네트워크 .....	27
2.4 혼합 네트워크의 허브 네트워크 추출 .....	28
2.4.1 네트워크 이산 계층화 .....	28
2.4.2 혼합 네트워크 임계치 .....	30
2.5 이산 계층화 검증 .....	33
2.6 소 결 .....	36
참고문헌 .....	38
제 3 장 허브 네트워크의 취약성 분석 .....	41
3.1 서 론 .....	41

3.2 허브 네트워크 취약성 분석 .....	43
3.3 네트워크 취약성 분석을 이용한 확산 해석 .....	45
3.3.1 네트워크 허브와 확산 .....	46
3.3.2 모델의 적용 .....	51
3.4 소 결 .....	68
참고문헌 .....	71
<b>제 4 장 확률을 이용한 허브 네트워크 군집화 분석 .....</b>	<b>75</b>
4.1 서 론 .....	75
4.2 확률 재규격화를 이용한 군집화 .....	77
4.2.1 확률 재규격화 .....	77
4.2.2 군집화 모델 개발 .....	79
4.3 군집화 모델의 구현 및 고찰 .....	86
4.3.1 CPC 모델의 구현 .....	86
4.3.2 CPC 모델의 검증 .....	93
4.3.3 기존 모델과 비교 .....	95
4.4 소 결 .....	97
참고문헌 .....	99
<b>제 5 장 허브 네트워크 추출과 분석 모델의 적용 .....</b>	<b>103</b>
5.1 연속체 구조물의 응력 흐름에 대한 군집화 적용 .....	103
5.1.1 연속체 구조물의 이산화 .....	103
5.1.2 허브 네트워크 추출과 진화론적 최적화 구조 기준 .....	105
5.1.3 응력 흐름의 군집화 분석 .....	110
5.2 뼈대 구조의 파괴 모드 해석 .....	115
5.2.1 뼈대 구조에서 CPC 구현 .....	117
5.2.2 뼈대 구조의 네트워크 치환 .....	120
5.2.3 파괴 모드 해석 .....	121
5.3 확산 네트워크의 추출과 취약성 분석 .....	125
5.3.1 대상 자료 .....	125
5.3.2 확산 네트워크의 추출 .....	127
5.3.3 확산 네트워크의 취약성 분석 .....	128

참고문헌 .....	132
<b>제 6 장 결 론 .....</b>	<b>135</b>
6.1 요약 및 결론 .....	135
6.2 향후 연구 .....	138
<b>Abstract .....</b>	<b>141</b>



## 표 목 차

Table 1.1 The outline of Study .....	12
Table 3.1 Centrality in HPAI diffusion network .....	56
Table 3.2 Centrality infected/suspected farms .....	58
Table 3.3 Factor analysis using a binary logistic model .....	66
Table 3.4 Outbreak prevention index with quarantined farm ·	68
Table 4.1 Distribution of the node in the example of the network .....	87
Table 4.2 Distribution of the Resistance in the node .....	87
Table 4.3 Explanation to the probability of stress flow in node .....	89
Table 4.4 Probability of stress flow in node .....	90
Table 4.5 Comparison with Monte-Carlo simulation .....	94
Table 4.6 Comparison with the existing models .....	95
Table 5.1 Boundary Condition in portal frame .....	116
Table 5.2 Comparison of failure probability in CPC and the existing method .....	122



## 그림 목차

Fig. 2.1 Cosmic microwave background radiation, white noise and unit hydrograph .....	20
Fig. 2.2 Random network and Scale-free network .....	21
Fig. 2.3 Comparison with Weibull distribution and Gamma distribution .....	23
Fig. 2.4 Gamma distribution with scale and shape factor .....	24
Fig. 2.5 Random network and connectivity distribution .....	25
Fig. 2.6 Scale-free network and connectivity distribution .....	26
Fig. 2.7 Real-world network and connectivity distribution .....	27
Fig. 2.8 Extraction of Hub network in Real world connection	29
Fig. 2.9 Traffic network and connectivity distribution .....	33
Fig. 2.10 Main line network and connectivity distribution .....	34
Fig. 2.11 Branch line network and link distribution .....	35
Fig. 3.1 Location of 39 farms in Gimje, South Korea .....	48
Fig. 3.2 Link distribution on poultry related business network ..	49
Fig. 3.3 Data collection for consideration of direct spread by poultry-related business vehicles or personnel visiting farms and indirect spread by transmit through air .....	50
Fig. 3.4 HPAI spread network with poultry related business data .....	51
Fig. 3.5 Flow chart of the study procedure .....	53
Fig. 3.6 Centrality analysis results from HPAI spread network .....	57

Fig. 3.7 The direct HPAI spread network	
with two factors, animal feed and medicine .....	60
Fig. 3.8 Trace the infection route diagram .....	61
Fig. 3.9 System flow of constructing the infection route	
diagram based on adjacent matrix data .....	63
Fig. 3.10 Comparison of the tendency of HPAI spread	
with/without quarantining farm .....	67
Fig. 4.1 Convolution in network	
that calculate link and node .....	78
Fig. 4.2 Research scope of directed, hierarchy hub network ..	80
Fig. 4.3 Convolution in directed, hierarchy hub network .....	85
Fig. 4.4 System flow of Clustering	
with Probabilistic Convolution model .....	94
Fig. 4.5 Example of the network for CPC model .....	86
Fig. 4.6 Analysis process of the example .....	88
Fig. 4.7 Result of the stress flow using CPC model .....	91
Fig. 4.8 Result of the stress flow using CPC model	
with the criteria of clustering .....	92
Fig. 4.9 Comparison of Number of iteration	
with the existing method .....	97
Fig. 5.1 Each truss element with coordinates .....	105
Fig. 5.2 Simple beam with distributed load .....	108
Fig. 5.3 Stress path in simple beam	
with maximum value of performance index .....	108
Fig. 5.4 Connectivity distribution in truss network .....	109
Fig. 5.5 Performance index with removal ratio .....	110
Fig. 5.6 Load of node in stress path model .....	112
Fig. 5.7 System flow of CPC in stress path .....	113

Fig. 5.8 Stress flow using CPC model with $W_{positive} = 30\%$ .....	114
Fig. 5.9 System flow using CPC model with $W_{positive} = 70\%$ ..	114
Fig. 5.10 Theoretical result(Nawy, E. G., 1990) .....	114
Fig. 5.11 Example of portal frame with concentrated load .....	116
Fig. 5.12 System flow of CPC in portal frame .....	117
Fig. 5.13 Mechanism of stress path in portal frame .....	118
Fig. 5.14 Set of Fundamental Mechanisms with geometry, loading and potential yield hinges ..	119
Fig. 5.15 Network model for the portal frame .....	120
Fig. 5.16 Stress path of portal frame using CPC model .....	123
Fig. 5.17 Comparison with MCS and CPC in the number of nodes in portal frame network .....	124
Fig. 5.18 FMD network and structure of FMD network .....	126
Fig. 5.19 Connectivity distribution in FMD network .....	128
Fig. 5.20 Extraction of hub network from FMD network and Connectivity distribution in hub network .....	129
Fig. 5.21 Ring Vaccination in random network from FMD network .....	131



# 제 1 장 서 론

## 1.1 연구배경

미소요소가 복잡하게 얽혀있는 시스템을 모의할 때 네트워크는 편리한 도구가 될 수 있다. 요소를 절점으로, 요소간의 관계를 연결선으로 구분하여 표시함으로써 복잡한 시스템을 수학적으로 나타낼 수 있어 표현과 해석이 용이하기 때문이다. 이런 이유로 실세계를 네트워크로 표현하는 연구는 다양한 분야에서 시도된 바 있다 (Albert et al., 1999; Jolly et al., 2001; Newman et al., 2002).

Barabasi 등은 네트워크의 절점과 연결선의 관계를 연결선 수 분포로 나타냈는데, 이는 종전의 균등하게 설정된 격자 네트워크를 이용하던 방식을 불균등 네트워크로 전환하는 계기가 되었다 (Barabási and Albert, 1999). 균등 격자 네트워크의 연결선 수 분포는 한 절점이 4개 정도의 연결선을 가지므로 평균값이 4인 정규분포와 같은 형태가 된다. 그러나 실세계에서는 절점의 중요도에 따라 연결선 수가 다르게 된다. 특히 중요도가 높은 절점은 많은 연결선 수를 가지게 되므로 허브라고 부르며, 이 허브는 네트워크를 정의하고 규정짓는 성질을 가진다. 네트워크의 특성을 결정하는 허브 절점은 다른 절점과 비교해서 연결선을 독점하고 있기 때문에 허브 네트워크의 연결선 수 분포는 매우 불균등하다.

복잡계에 대한 연구로 네트워크가 주목받기 시작한 것은 실제 세계를 네트워크의 불균등한 구조로 표현하기 시작할 때부터이다 (Barabási et al., 1999; Redner, 1998; Jeong et al., 2000; Venter et al., 2001). 이 일련의 연구들을 통해서 자연계는 균등한 정규분포보다 불균등한 멱급수 분포와 더 유사한 형태라는 것이 증명되었다. 즉, 네트워크는 복잡계의

발현 현상을 해석하기 위한 중요한 도구이며, 그 중에서 허브 절점으로 이루어진 허브 네트워크는 네트워크의 특징을 잘 표현할 수 있으므로 네트워크 연구의 핵심이 되고 있다.

네트워크의 불균등성을 나타내는 허브 네트워크에서 허브 절점간의 중요도에 따라 위상차를 분석하는 연구는 네트워크 구조의 견고성 및 취약성을 분석하는데 다양하게 이용된다. 허브 절점은 수많은 절점들에게 영향을 미치기 때문에 파괴되면 다른 절점들의 연쇄적 파멸로 전체 구조가 쉽게 무너질 수 있기 때문에 구조적 안정성을 위해 가장 우선적으로 보호해야 할 절점이 된다. 또, 네트워크의 확장은 허브 네트워크의 연결선 확장과 직결되어 있으므로 네트워크의 수축, 팽창과 같은 외적 변화를 판단하는 요소가 된다. 따라서 허브 네트워크에서 허브 절점을 찾고 이를 통해서 전체 구조를 해석하는 방법은 복잡계 현상을 이해할 수 있는 중요한 분석 가운데 하나이다.

허브 네트워크의 임의의 절점은 일부 절점과는 강하게 연결되어 있으며 다른 절점과는 약하게 연결되어 있다. 연결성의 강도에 따라 강하게 연결된 절점들을 모으면 작은 집단이 형성될 수 있다. 같은 집단에 소속된 절점들은 유사한 특성을 가지기 때문에 서로의 특성을 예측할 수도 있고 절점 간의 그룹화를 통해서 전체 구조의 단순화도 가능하기 때문에 네트워크 군집화는 네트워크 내부의 관계를 파악하는 분석이 될 수 있다. 그러나 절점과 연결선이 확률적 특성을 가지게 되면 대상을 더 정확하게 모의할 수 있으므로 확률과 결합된 새로운 확률 군집화 방법이 필요하다.

본 연구는 실세계에서는 정규분포와 멱급수 분포가 적절하게 혼재되어 있다는 현상 분석에서 시작되었다. 주요 허브 절점을 중심으로 연결된 네트워크는 수학적으로 멱급수 분포로 표현되고, 격자망과 같이 균등한 네트워크는 수학적으로 정규분포로 표현된다. 앞에서 서술한 대로 만약 멱급수 분포와 정규분포가 혼재된 문제에서 네트워크의 특성을 나타내는

허브 네트워크, 즉 멱급수 분포를 분리한다면 네트워크 문제에 효율적이고 신속한 해를 얻는데 도움이 될 것이다.

멱급수 분포와 정규분포가 혼합된 분포는 웨이블 분포 또는 감마 분포를 이용해서 단일 분포 형태로 나타낼 수 있다. 이런 특성은 두 분포를 두 모수로 대체할 수 있기 때문에 계산량을 줄이고 특성을 간단하게 표현할 수 있다.

따라서 본 연구에서는 먼저 감마분포를 이용하여 허브 네트워크를 추출하는 방법론을 정립하고 추출된 허브 네트워크를 이용하여 네트워크의 상생, 소멸 과정을 모사하고자 한다. 다음으로 허브 네트워크와 확률로 결합된 확률 군집화 방법을 개발하여 네트워크 내부의 중요도를 분석하였다. 마지막으로 정립된 방법을 지역시스템공학에 자주 발생하는 확산 문제, 구조 신뢰성 문제에 적용함으로써 유용성을 검증할 것이다.

## 1.2 연구사

### 1.2.1 복잡계 네트워크

Euler는 문제 해결을 위한 수단으로 그래프를 처음으로 이용하였으며 (Euler, 1741), 이후 그래프는 경로 문제 (Dijkstra, 1959; Ford and Fulkerson, 1956), 위상 문제 (Freeman, 1979; Beauchamp, 1965) 등을 해결하는데 이용되었다. Erdos와 Reny는 그래프의 연결선과 절점의 관계를 분석하여 연결선 수 분포의 관계가 가우시안 분포와 동일한 무작위 네트워크(Random network)를 발견하였다 (Erdős and Rényi, 1959). 무작위 네트워크는 가우시안 분포를 이용해서 실세계를 모사하고자 하였다.

네트워크 이론이 복잡계 문제에 본격적으로 적용되기 시작한 것은

Albert가 월드 와이드 웹의 네트워크 구조가 멱급수의 법칙으로 이루어져 있음을 보였을 때부터이다 (Albert et al., 1999). 이 연구는 20~30만 개의 웹사이트 연결 관계를 분석한 결과 무작위로 균등하게 연결된 가우시안 분포가 아닌 일부 웹사이트에 연결성이 집중되어 있는 멱급수의 법칙을 따르는 분포를 발견하였다. Huberman과 Kumar는 웹 네트워크의 구조가 불균등한 분포 형태를 가지는 것을 보다 많은 자료를 가지고 확인했으며 (Huberman and Adamic, 1999; Kleinberg et al., 1999). 인맥관계, 논문 인용, 유전자 구조 등 실세계의 다양한 문제에서 불균등한 멱급수의 분포를 가지는 네트워크로 모델링할 수 있음이 연구되었다 (Barabási et al., 1999; Redner, 1998; Jeong et al., 2000; Venter et al., 2001). Barabasi는 멱급수 분포를 가지는 불균등한 구조의 네트워크를 척도 없는 네트워크 (Scale-free network) 즉, 허브 네트워크라고 정의하고 그 생성과 진화에 대해서 소개하였다 (Barabási and Albert, 1999). 임의로 생성되는 절점이 연결될 확률은 다른 절점들이 가지고 있는 연결선 수에 비례하여 결정되는 선호적 연결(preferential attachment)의 법칙을 따른다는 것을 보였다. 무작위 네트워크와 다른 특성을 가지는 허브 네트워크는 연결선을 일부 허브 절점이 독점하고 있으며 복잡계를 대표하는 특성을 가지고 있기 때문에 다양한 대상을 모델링하기 위한 도구로 이용되었다 (Pastor-Satorras and Vespignani, 2001; Caldarelli, 2007; Stauffer et al., 2003).

그러나 실제 문제는 순수하게 허브 네트워크만으로 구성되어 있지 않고 무작위 네트워크가 혼재되어 있기 때문에 허브 네트워크의 해석 방법을 그대로 적용하면 해석의 효율성과 결과의 정확도가 낮아질 수 있다. 따라서 본 연구에서는 복잡계 해석의 효율성과 정확도를 향상시키는 방법에 대해서 고찰하고자 하였다.

## 1.2.2 네트워크 추출

네트워크는 적용하고자 하는 대상의 절점이 매우 많으며 연결선이 복잡하게 얽혀있기 때문에 효율성과 정확성을 향상시키기 위한 연구는 매우 중요하다. 기본적으로 네트워크 해석을 위해서는 모든 절점 간의 관계를 연산해야하기 때문에 알고리즘 개선에 관한 연구가 다수 진행되었다 (Israeli and Wood, 2002; Nardelli et al., 2003; Newman, 2004). 그러나 절점이나 연결선의 수가 줄어들지 않는 이상 해석의 속도를 향상시키는데 한계가 있다. 그리고 중심성 분석, 군집화 분석과 같은 네트워크 해석 결과의 정확성을 향상시키기 위한 방법들이 제안되고 있지만 (Girvan and Newman, 2002; Newman, 2004; Reichardt and Bornholdt, 2004) 크게 주목 받지 못하고 있다 (Liu et al. 2005). Boccaletti 등은 네트워크 해석의 정확도를 향상시키기 위해서는 방법론의 개선보다 전체 구조의 특성에 영향을 미치지 못하는 불필요한 절점들의 간섭을 우선적으로 제거할 필요가 있다고 지적하였다 (Boccaletti et al., 2006). 따라서 해석의 효율성과 정확성을 동시에 향상시킬 수 있는 방법은 대표성을 가지는 네트워크를 추출하는 전처리 과정을 활용할 필요가 있다.

서로 다른 차원의 네트워크가 혼재되어 있는 상태에서 네트워크를 분류하는 시도는 다양하게 진행되었다. Sema 등은 절점의 특성이 다른 네트워크를 대상으로 혼재되어 있는 상태에서 네트워크를 분석하고 그 결과에서 특성이 다른 절점을 분류하고자 하였다 (Sema et al., 2011). 이 방법은 전처리 과정으로 적용된 것이 아니기 때문에 해석의 효율성과 정확성을 향상시키는데 도움이 되지 못한다. Barthélemy 등은 연결선의 속성이 다른 확산을 모의하는 허브 네트워크를 대상으로 매개체의 확산 속도에 따라서 계층을 분류하는 방법을 개발하였고 (Barthélemy et al., 2004) Olaf는 인간 뇌를 다양한 계층으로 구분된 네트워크로 분류하는 방법을 소개하였다 (Sporns, 2011).

그러나 이 방법들은 각 사례에 적용된 절점과 연결선의 특성에 따라서 분류 기준을 정의하였기 때문에 범용적인 방법에는 적용할 수 없는 한계를 가진다. 본 연구의 2장에서는 네트워크의 가장 기본적인 특성인 연결선 수 분포를 이용해서 혼합 네트워크의 특성을 대표하는 허브 네트워크를 추출하는 기준을 개발하였다.

### 1.2.3 중심성 분석

허브 네트워크에는 무작위 네트워크와 다르게 다른 절점과 비교해서 연결선을 독점하고 있는 허브 절점들이 존재한다. 이 허브 절점들은 네트워크 구조의 변화에 중요한 역할을 한다. Albert 등은 허브 절점들이 외부에서 공격받을 때와 그렇지 않을 때의 네트워크 구조의 변화를 비교했다 (Albert et al., 2000). 허브 절점을 가진 네트워크는 외부에서 공격받을 때 전체 구조가 붕괴될 위험을 가진다. Jeong 등은 신진대사 네트워크에서 허브 절점이 집중적으로 공격할 때 세포 및 생명체는 심각한 타격을 받을 수 있음을 밝혔다 (Jeong et al., 2000). 유클라디안 공간에서 무작위 네트워크의 절점들이 균등하게 3~4개 정도의 연결선을 가지는 반면에 허브 네트워크의 절점들은 매우 불균등한 분포를 가지기 때문에 전체 구조의 붕괴까지 발전할 수 있다. Sachtjen 등은 전력망 네트워크를 대상으로 일부 절점에 집중된 전력량이 문제를 일으키며 전체 전력망이 파괴된 사례를 분석했다 (Sachtjen et al., 2000).

허브 네트워크에서 연쇄적 과열현상은 확산 현상의 모의와 밀접한 연관이 있다. 절점과 절점 간의 연결선에 따라서 확산 매개체가 이동되기 때문에 많은 절점에게 동시에 영향을 주는 허브 절점들을 탐색하는 것은 확산 차단을 위한 가장 효과적인 방법이다. Jolly 등은 상관관계에 대해서 네트워크를 이용하여 주요 절점을 분석하고 (Jolly et

al., 2001) Newman 등은 이메일의 연결 관계를 분석한 확산 네트워크를 구축하고 허브 절점들을 기반으로 바이러스의 확산 양상을 분석하였다 (Newman et al., 2002). 그리고 절점들의 상태를 감염상태, 회복상태 등으로 구분하고 관계 기작을 감염율, 회복율을 이용해서 정의한 SIS, SIR 모델 네트워크를 구축하고 확산 양상을 분석한 연구도 있다 (Eames and Keeling, 2002).

이 연구들은 확산을 모의하는 네트워크에서 허브 절점들을 이용하여 효율적으로 확산을 차단하는 방안을 제시하였다. 본 연구의 3장에서는 국내 HPAI 발병 사례를 대상으로 확산 매개체를 분석하고 이를 네트워크로 구축하여 허브 절점을 탐색하고 연쇄 현상을 모의하며 가축전염병의 확산을 방제하기 위한 방안을 제안하였다.

#### 1.2.4 군집화 분석

불규칙한 구조의 네트워크를 몇 개의 소집단으로 분류하는 군집화 방법은 각 절점들의 특성을 분석하고 전체 구조를 단순화시킬 수 있다 (Newman, M. E., 2004.). Girvan과 Newman은 각 절점간의 연결선의 밀도를 분석하여 밀도가 높은 연결선을 점차 제거해나가면서 소집단으로 구분하는 방법을 개발하였다 (Girvan and Newman, 2002). 이 방법은 소집단 내 절점들의 연결성은 높지만 소집단과 소집단 간의 연결성은 높지 않다는 착안에서 비롯되었다. 그러나 단계마다 절점 간 연결성의 밀도를 산정해야하기 때문에 절점 수가 많은 네트워크에는 적합하지 않았으며 해석의 효율성을 개선하기 위한 연구가 시도되기도 하였다 (Gudkov and Nussinov, 2002). Newman은 연결선을 없애는 방법과 다르게 유사한 특성을 가지는 절점들을 연결하면서 군집화 하는 방법으로 기존 방법의 효율성을

개선하고자 하였다 (Newman, 2004.). 이외에 계층형 네트워크에서 모든 절점들의 관계를 정의하고 하층집단에서부터 특성이 유사한 절점들을 묶어가면서 점진적으로 소집단을 만들어내는 방법도 있다 (Reichardt and Bornholdt, 2004).

그러나 이 방법들은 절점과 연결선이 결정론적 값을 가지는 네트워크를 대상으로 한 것이며 확률론적 범위를 가지는 네트워크를 대상으로 적용하는데 한계가 있다. Friedman은 연결선의 값을 조건부 확률로 하고 군집화 하는 방법을 생물학 네트워크에 적용하였으나 (Friedman, 2004.) 절점 간의 관계 및 계층이 이미 정의되어 있는 분자 생물학이 아닌 범용적인 사례에 적용하는데 한계가 있다. Leskovec 등과 Peter 등은 절점이 확률적 특성을 가질 때 군집화 하는 방법을 개발하였으나 연결선과 절점이 모두 확률적 값을 가지는 네트워크를 대상으로 해석하기는 어렵다 (Leskovec et al., 2009; Mucha et al., 2010). 따라서 본 연구의 4장에서는 연결선과 절점이 모두 확률적 범위를 가지며 서로 상응하는 조건을 가지고 있는 네트워크에서 컨볼루션이라는 확률적 연산을 이용해서 군집화 하는 방법을 개발하고 적용하였다.

## 1.3 연구목적 및 범위

### 1.3.1 감마분포를 이용한 허브 네트워크 추출

실세계는 무작위로 연결된 네트워크와 전체 구조를 대표할 수 있는 허브 네트워크가 혼합되어 있다. 전체 구조를 대표할 수 있는 허브 네트워크를 추출할 수 있다면 효율적으로 실세계에서 적용할 수 있는 문제를 해석할 수 있다.

절점과 연결선으로 구성되는 네트워크에서 둘의 관계에 대한 분포는 가장 중요한 특성이 될 수 있다. 무작위 네트워크와 허브 네트워크는 연결선 수 분포에 있어서 각각 정규 분포와 멱급수 분포를 가진다. 두 분포는 감마분포로 나타내면 형상계수와 규모계수를 통해서 규격화될 수 있고 두 개의 지수를 이용하면 혼합되어 있는 네트워크에서 허브 네트워크를 추출할 수 있는 기준을 산정할 수 있다.

본 연구에서는 서로 다른 두 네트워크의 연결선 수 분포를 포괄적으로 분석하여 혼합 네트워크에서 허브 네트워크를 추출하는 방법론을 정립하였다. 그리고 이 방법론을 간선과 지선이 혼재된 교통 네트워크에 적용하여 검증하고자 하였다. 혼합 네트워크에서 허브 네트워크를 추출하는 방법은 다양한 문제에서 네트워크로 모델링하고 해석할 때 전처리 과정으로 유용하게 활용될 수 있을 것으로 기대된다.

### 1.3.2 허브 네트워크의 취약성 분석

무작위 네트워크와 달리 허브 네트워크에서는 절점들의 중요도가 다르다. 연결선 수 분포의 불균등성 때문에 일부 허브 절점들이 대다수의 연결선을 독점하고 있다. 따라서 허브 절점들은 네트워크 구조에서 다른 절점과 비교해서 상대적으로 중요한 역할을 가진다. 전체 절점 중 허브 절점은 극소수이기 때문에 네트워크가 견고성을 가질 수 있으나 외부로부터 집중적으로 허브 절점이 공격받는다면 연쇄적으로 다른 수많은 절점들에게 영향을 줘서 전체 구조가 쉽게 무너질 수 있는 취약성도 동시에 가지고 있다.

본 연구에서는 허브 네트워크를 대상으로 허브 절점을 탐색하는 방법론에 대해서 고찰하고 실제 확산 문제를 모의할 수 있는 네트워크에 적용하였다. 허브 절점은 네트워크에서 주요한 지점에 위치하고 있고 수많은 절점에 영향력을 미치기 때문에 허브 절점 탐색은 확산 문제를 해석

할 때 가장 중요한 분석 가운데 하나이다. 허브 절점을 이용해서 확산 양상을 분석하고 차단할 수 있는 대책을 제안하였다.

### 1.3.3 확률을 이용한 허브 네트워크 군집화 분석

불균등한 분포를 가지는 허브 네트워크에서 유사한 특성을 가진 절점들의 집단을 찾아내는 군집화 방법은 절점들의 특성을 분석하는데 유용하게 활용된다. 그러나 절점과 연결선이 모두 확률적 분포를 가지는 네트워크를 대상으로 하는 군집화 연구는 확률적 연산의 어려움으로 포괄적인 분석을 수행하는데 한계를 가진다.

본 연구에서는 절점과 연결선의 확률적 특성을 동시에 고려할 수 있으며 두 개의 확률적 분포를 연산하는 방법으로 컨볼루션이라는 수학적 연산자를 이용하였다. 이를 통해서 절점과 연결선이 모두 확률적 분포를 가지는 네트워크를 대상으로 두 개의 확률 분포를 새로운 확률 분포로 규격화시키고 그 기준에 따라 군집화가 가능한 모델을 개발하였다. 개발된 모델은 몬테카를로 시뮬레이션을 이용한 수치적 해법과의 비교를 통해서 정확도를 검증하였으며 기존 모델과 해석의 효율성 등을 비교하여 활용성을 검토하였다.

## 1.4 연구구성

본 연구는 총 6장으로 구성되어 있으며 각 장은 다음과 같은 내용으로 구성되어 있다.

제 1 장은 본 논문의 연구목적과 각 장별 연구를 요약하여 설명하고 있다.

제 2 장에서는 네트워크 해석의 전처리 과정으로 혼합 네트워크를 대상으로 감마분포 시뮬레이션을 이용해서 허브 네트워크를 추출하는 방법론을 정립하였다.

제 3 장은 혼합 네트워크에서 추출된 허브 네트워크를 대상으로 네트워크 허브 절점 탐색 방법론을 고찰하고 실제 가축전염병의 확산 문제에 적용하였다.

제 4 장은 고정된 경계조건에서 네트워크 내의 요소 분석을 위해 결정론적 값이 아닌 확률론적 범위를 가진 절점과 연결선으로 구성된 허브 네트워크를 대상으로 컨볼루션을 이용하여 군집화 하는 방법론을 개발하고 기존의 방법과 비교 고찰하였다.

제 5 장에서는 각 장에서 소개된 방법론을 연계하여 구조 신뢰성과 확산 문제에 적용하고자 하였다. 추출 방법론과 군집화 분석을 이용하여 연속체를 이산화한 구조와 뼈대 구조의 파괴확률을 해석하였고 추출 방법론과 취약성 분석을 단계적으로 수행하여 확산 문제를 해석함으로써 본 논문에서 정립한 방법론의 활용성에 대하여 고찰하였다.

제 6 장에서는 전체적인 연구의 내용과 결과를 종합적으로 정리하였으며 연구의 성과와 한계를 통해 향후 연구과제도 함께 제시하였다.

Table 1.1 The outline of Study

	<b>Extraction of hub network</b>	<b>Vulnerability analysis</b>	<b>Detecting community</b>
<b>Study Area</b>	Real-world network	Undirected Hub network in Flexible Boundary Condition	Directed Hub network in Fixed Boundary Condition
	<b>Complex network</b>		
<b>Method</b>	Gamma distribution	Centrality analysis	Convolution
	Network analysis		
<b>Conclusion</b>	Classification of real-world network	Macroscopic analysis of network	Microscopic analysis of network
<b>Application</b>	Pre-Processing in network analysis	Analysis of spreading in network	Structural analysis with network

## 참고문헌

1. Albert, R., H. Jeong and A.-L. Barabási, 1999. Internet: Diameter of the world-wide web. *Nature* 401.6749 : 130-131.
2. Albert, R., H. Jeong and A.-L. Barabási, 2000, Error and attack tolerance of complex networks. *Nature* 406.6794 : 378-382.
3. Barabási, A.-L., R. Albert and H. Jeong. 1999. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications* 272(1): 173-187.
4. Barabási, A.-L. and R. Albert, 1999. Emergence of scaling in random networks. *science* 286.5439 : 509-512.
5. Barabási, A.-L., R. Albert and H. Jeong, 1999. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications* 272.1 : 173-187.
6. Barthélemy, M., A. R. Barrat, Pastor-Satorras and A. Vespignani, 2004. Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. *Physical Review Letters* 92.17 : 178701.
7. Beauchamp, M. A., 1965. An improved index of centrality. *Behavioral Science* 10.2 : 161-163.
8. Boccaletti, S., V. Latora, Y. Moreno, M. Chavez and D. U. Hwang, 2006. Complex networks: Structure and dynamics. *Physics reports* 424.4 : 175-308.
9. Caldarelli, G., 2007. Scale-Free Networks: complex webs in nature and technology. OUP Catalogue.
10. Dijkstra, E. W., 1959. A note on two problems in connexion with graphs. *Numerische mathematik* 1.1 : 269-271.
11. Eames, K. T., and M. J. Keeling, 2002. Modeling dynamic and

- network heterogeneities in the spread of sexually transmitted diseases. *Proceedings of the National Academy of Sciences* 99.20 : 13330–13335.
12. Erdős, P. and A. Rényi, 1959. On random graphs. *Publicationes Mathematicae Debrecen* 6 : 290–297.
  13. Euler, L., 1741. *Solutio problematis ad geometriam situs pertinentis*. *Commentarii academiae scientiarum Petropolitanae* 8 : 128–140.
  14. Ford, L. R. and D. R. Fulkerson, 1956. Maximal flow through a network. *Canadian Journal of Mathematics* 8.3 : 399–404.
  15. Freeman, L. C., 1979. Centrality in social networks conceptual clarification. *Social networks* 1.3 : 215–239.
  16. Friedman, N, 2004. Inferring cellular networks using probabilistic graphical models. *Science* 303.5659 : 799–805.
  17. Huberman, B. A. and L. A. Adamic, 1999. Internet: growth dynamics of the world-wide web. *Nature* 401.6749 : 131–131.
  18. Girvan, M. and M. E. Newman, 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99.12 : 7821–7826.
  19. Gudkov, V. and S. Nussinov, 2002. Graph equivalence and characterization via a continuous evolution of a physical analog. *arXiv preprint cond-mat/0209112*.
  20. Israeli, E. and R. K. Wood, 2002. Shortest path network interdiction. *Networks* 40.2 : 97–111.
  21. Jeong, H., B. Tombor, R. Albert, Z. N. Oltvai and A.-L. Barabási, 2000. The large-scale organization of metabolic networks. *Nature* 407.6804 : 651–654.
  22. Jolly, A. M., S. Q. Muth, J. L. Wylie and J. J. Potterat, 2001.

- Sexual networks and sexually transmitted infections: a tale of two cities. *Journal of Urban Health* 78.3 : 433–445.
23. Kleinberg, J. M., R. Kumar, P. Raghavan, S. Rajagopalan and A. S. Tomkins, 1999. The web as a graph: Measurements, models, and methods. *Computing and combinatorics*. Springer Berlin Heidelberg 1–17.
  24. Leskovec, J., K. J. Lang, A. Dasgupta and M. W. Mahoney, 2009. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics* 6.1 : 29–123.
  25. Mucha, P. J., T. Richardson, K. Macon, M. A. Porter and J. P. Onnela, 2010. Community structure in time-dependent, multiscale, and multiplex networks. *Science* 328.5980 : 876–878.
  26. Nardelli, E., G. Proietti and P. Widmayer, 2003. Finding the most vital node of a shortest path. *Theoretical computer science* 296.1 : 167–177.
  27. Newman, M. E. J., 2003. The structure and function of complex networks. *SIAM review* 45(2): 167–256.
  28. Newman, M. E., F. Stephanie and J. Balthrop, 2002. Email networks and the spread of computer viruses. *Physical Review E* 66.3 : 035101.
  29. Newman, M. E., 2004. Fast algorithm for detecting community structure in networks. *Physical review E* 69.6 : 066133.
  30. Pastor-Satorras, R. and A. Vespignani, 2001 Epidemic spreading in scale-free networks. *Physical review letters* 86.14 : 3200.
  31. Redner, S., 1998. How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems* 4.2 : 131–134.

32. Reichardt, J. and S. Bornholdt, 2004. Detecting fuzzy community structures in complex networks with a Potts model. *Physical Review Letters* 93.21 : 218701.
33. Sachtjen, M. L., B. A. Carreras. and V. E. Lynch, 2000. Disturbances in a power transmission system. *Physical Review E* 61.5 : 4877.
34. Sema, N., L. Matthews, P.R. Bessell, S. WJ Reid and R.R. Kao, 2011. Generating social network data using partially described networks: an example informing avian influenza control in the British poultry industry. *BMC Veterinary Research* 7(66).
35. Sporns, O. 2011. The human connectome: a complex network. *Annals of the New York Academy of Sciences* 1224.1 : 109-125.
36. Stauffer, D., A. Aharony, C. L. da Fontoura and J. Adler, 2003. Efficient Hopfield pattern recognition on a scale-free neural network. *The European Physical Journal B-Condensed Matter and Complex Systems* 32.3 : 395-399.
37. Strogatz, S. H. 2001. Exploring complex networks. *Nature* 410: 268-276.
38. Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton and E. Beasley, 2001. The sequence of the human genome. *science* 291.5507 : 1304-1351.

## 제 2 장 감마분포를 이용한 허브 네트워크 추출

### 2.1 서론

네트워크는 절점과 연결선으로 구성되어 있기 때문에 둘의 관계를 나타내는 연결선 수 분포와 같은 고유의 특성을 가진다. 연결선 수의 분포에 따라서 네트워크는 크게 가우시안 분포를 가지는 무작위 네트워크(Random network)와 멱급수 분포를 가지는 허브 네트워크(Scale-free network)로 구분된다 (Barabási and Albert, 1999).

무작위 네트워크는 주로 절점들을 기하학적인 관계로 모사하며 허브 네트워크는 절점들의 위상학적 관계를 해석하고자 할 때 유용한 도구로 이용된다 (Barabási and Albert, 1999; Barabási et al., 2000; Holger et al., 2002; Wang and Chen, 2003). 기하학적인 기반의 무작위 네트워크는 절점이 가지는 연결선 수가 균등해서 모든 절점들의 중요도가 같기 때문에 네트워크 구조를 대표하는 허브 절점을 가지지 않는 반면에 허브 네트워크는 연결선이 거의 없는 수많은 절점들과 연결선을 독점하는 일부 허브 절점들로 불균등한 구조를 가진다. 두 개의 네트워크는 연결선 수 분포에 따라 특징적인 구조를 가지기 때문에 각각 다른 문제에 적용되는 경우가 많다 (Barabási and Albert, 1999; Romualdo and Alessandro, 2000; Boguna et al., 2003).

그러나 네트워크로 해석할 수 있는 실세계의 문제들은 무작위 네트워크와 허브 네트워크 가운데 완전히 하나의 특성을 따르기보다는 혼합되어 있는 경우가 다수 존재한다. 예를 들어 확산 문제에서 절점을 이동하는 매개체의 특성에 따라서 거리를 고려해야하는 무작위 네트워크와 거리를 고려할 필요가 없는 허브 네트워크 등이 여러 계층으로 혼합되어

있다. 인플루엔자의 전파는 자동차 등의 교통망을 통해서 전달이 되기도 하며 교통망보다 속도가 훨씬 빠른 항공망을 통해서도 전달된다. 인플루엔자가 교통망 네트워크를 통하면 도로에 따라서 전달되고 지리적 거리의 영향을 받지만 속도가 빠른 항공망 네트워크를 통하면 허브 항공을 중심으로 지리적 거리와 관계없이 전달된다. 서로 다른 특성을 가지는 네트워크가 혼합되어 있기 때문에 해석을 위해서는 무작위 네트워크와 허브 네트워크로 구분해서 모델링하고 해석할 필요가 있다. 혼합 네트워크에서 서로 다른 두 네트워크를 각각 추출할 수 있다면 본격적인 네트워크 해석의 전처리 과정으로 유용하게 활용될 수 있다.

특성이 다른 네트워크가 혼합되어 있는 문제를 해석하기 위한 연구는 다양한 사례에 적용되었다 (Albert, 2005; Barthelemy et al., 2004; Sema et al., 2011). 이 방법들은 절점들의 중요도와 연결선의 속성에 따라서 다른 특성의 네트워크를 분류하였다. 그러나 이는 범용적인 사례에 적용되는 경우가 아닌 특정한 사례에 적합한 해법이라는 한계를 가지며 절점 및 연결성의 중요도 분석을 위해 네트워크의 해석을 거쳐야하기 때문에 전처리 과정과는 차이가 있다. 만약 연결선 수 분포라는 네트워크의 고유한 특성을 이용하면 혼합 네트워크에서 허브 네트워크의 추출 기준을 정립할 수 있는 좋은 방안이 될 수 있다.

본 연구에서는 서로 다른 무작위 네트워크, 허브 네트워크의 연결선 수 분포를 비교하고 분석하였으며 이를 감마분포로 나타내고 형상계수와 규모계수로 분포를 규격화하였다. 그리고 고유의 형상계수와 규모계수를 이용해서 혼합 네트워크에서 허브 네트워크를 추출하기 위한 임계치를 산정하였다. 이 연구는 분포에 따른 네트워크의 특성을 정의할 수 있는 지표를 개발하고 그 지표를 이용해서 혼합 네트워크에서 대표성을 가지는 허브 네트워크를 추출하는 방법을 제시함으로써 네트워크를 활용하는 문제 해석의 효율성 및 해의 정확성을 향상시킬 수 있는 중요한 전처리 과정의 도구가 될 수 있을 것으로 기대된다.

## 2.2 혼합계의 네트워크 표현

실세계는 내용을 대표하는 계층과 배경을 나타내는 계층으로 구분된다. 관측 가능한 우주에서 지속되는 우주배경복사(Cosmic microwave background radiation)는 모든 방향에서 같은 세기로 균일하게 오는 초단파 잡음과 1/10만 수준 대역 차이의 불균등한 초단파 잡음이 혼합되어 있다. Fig. 2.1 (a)의 1/10만 수준의 불균등한 초단파 잡음은 온도가 높은 지역과 낮은 지역의 크기와 온도를 비교하여 초기우주 인력의 세기를 알아내고 물질이 얼마나 빨리 쌓이게 되었는지 추론할 수 있다. 또 암흑물질과 암흑에너지의 우주 구성 비율 등을 추정할 수 있으며, 우주가 영원히 팽창할 것인지 아니면 팽창속도가 느려질지 아니면 빨라질지 등에 관한 정보를 예측할 수 있다. 그리고 지구 내 모든 형태의 전자 장비와 매개체 내부에는 이산적인 전자의 운동으로 야기되는 잡음, Fig. 2.1 (d)와 같은 백색소음(White noise)이 모든 범위의 주파수에 대해 균등하게 분포되어 있다. 모든 생활소음에 균등하게 분포하고 있기 때문에 생활소음의 특성을 대표하지 않는 배경과 같은 소음이다. Fig. 2.1 (e)는 유효수분곡선으로 단위 수문도를 나타낸 것이다. 이 곡선은 지하수로부터 하천으로 유출되는 흐름인 기저유출과 강우에 따른 홍수유출이 혼합되어 있다. 홍수유출은 전체 수문도에서 강우 대한 내용에 해당하는 부분으로 배경에 해당하는 기저유출과 구분해서 해석할 필요가 있다. 따라서 수문학에서는 홍수유출을 추출하기 위해서 감수곡선법, 선형분리법 등을 이용하기도 한다.

네트워크도 절점과 연결선의 관계에 따라서 내용과 배경의 네트워크로 구분된다. 멱급수 분포를 가지는 허브 네트워크는 특성 있는 허브 절점을 가지기 때문에 내용이라 할 수 있고 가우시안 분포를 가지는 무작위 네트워크는 모든 절점이 균등하게 연결선 수를 가지고 있는 배경이라 할 수 있다.

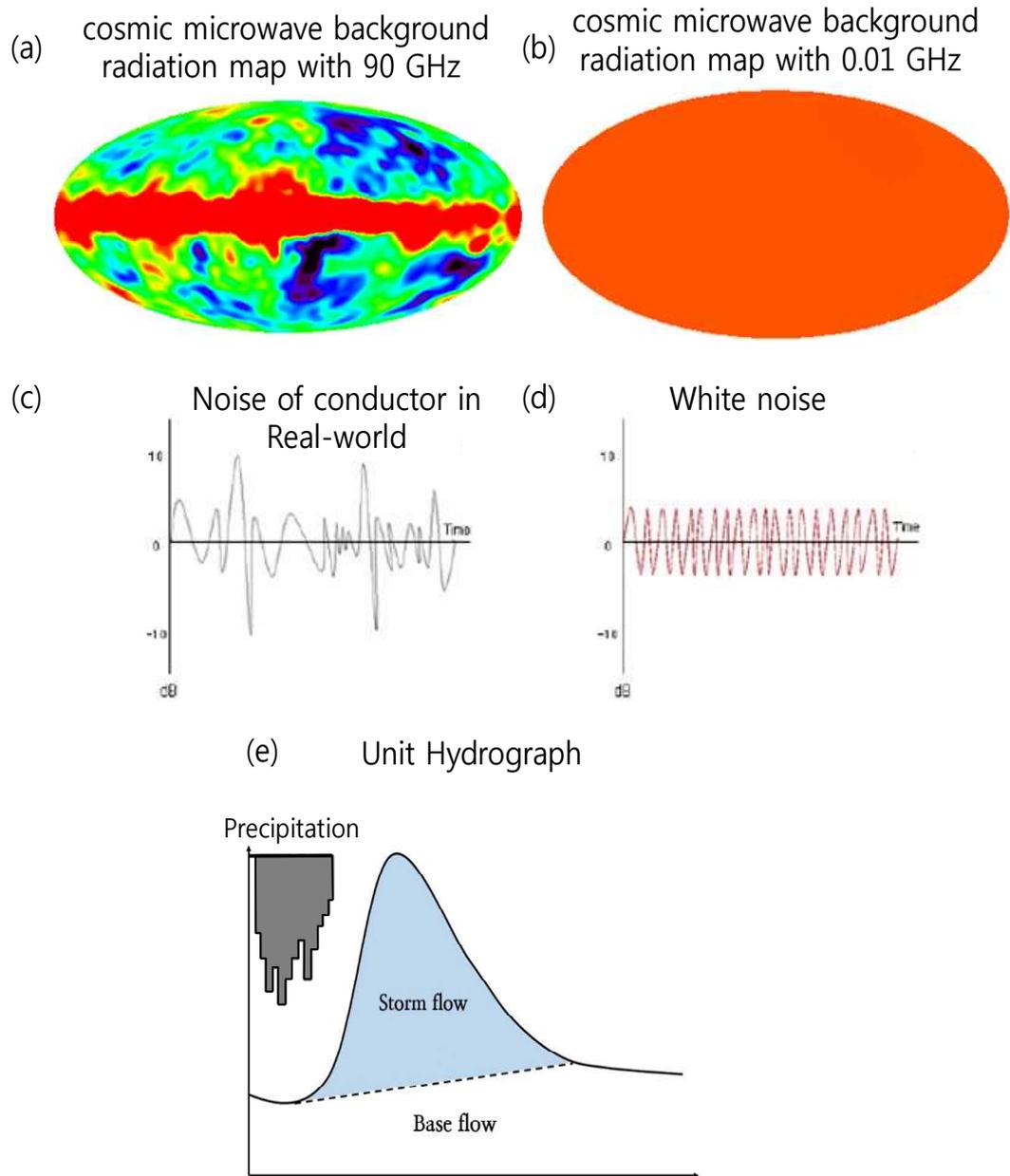


Fig. 2.1 Cosmic microwave background radiation, white noise and unit hydrograph

Fig. 2.2는 무작위 네트워크와 허브 네트워크의 연결선 분포와 대표적인 구조를 나타낸 것이다. 무작위 네트워크는 일부 범위에 절점들이 집중되어 있고 가우시안 분포의 형태를 보이며 연결선 수가 많은 허브 절점을 가지지 않는다. 반면에 허브 네트워크는 멱급수 분포에 따라서 다른 절점들과 달리 연결선 수가 집중된 허브 절점을 가지고 있다.

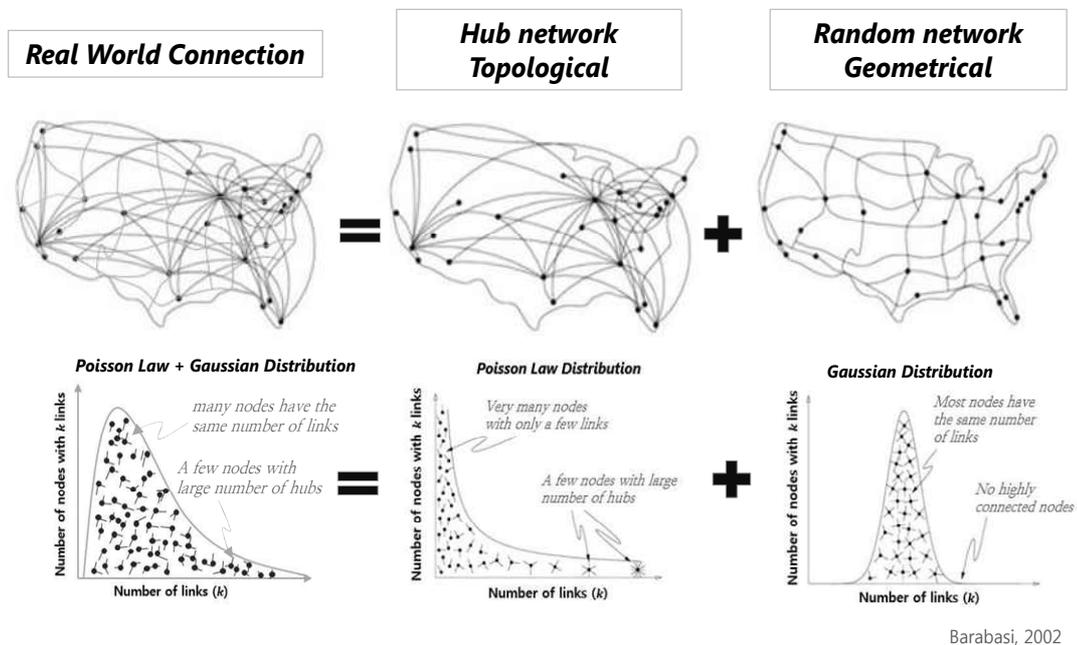


Fig. 2.2 Random network and Scale-free network  
(Barabási and Albert, 1999)

Fig 2.2의 무작위 네트워크를 미국의 도로망이라고 하고 허브 네트워크를 미국의 항공망이라고 한다면 도로망을 통해서 전달되는 매개체는 동부의 뉴욕에서 서부의 LA로 이동할 때 수많은 절점을 지나야하기 때문에 지리적 거리의 영향을 고려해야 한다. 반면에 항공망은 동부와 서부를 한 번에 연결할 수 있으며 속도도 도로망과 비교할 수 없을 정도로 빠르기 때문에 지리적 거리의 영향을 고려할 필요가 없다. 무작위 네트

워크와 허브 네트워크에서 연결선을 따라 매개체가 이동할 때는 각각 다르게 진행된다. 무작위 네트워크는 대부분의 절점이 가지는 연결선 수의 분포가 균등하기 때문에 매개체가 이동할 수 있는 절점 수의 차이가 거의 없다. 반면에 허브 네트워크에 특징적으로 존재하는 허브 절점에 도착하면 한 번에 많은 절점들에게 영향을 미치기 때문에 매개체가 이동할 수 있는 범위가 커진다. 균등성과 불균등성은 허브 네트워크와 무작위 네트워크의 구조를 구분하는 가장 중요한 특성이다.

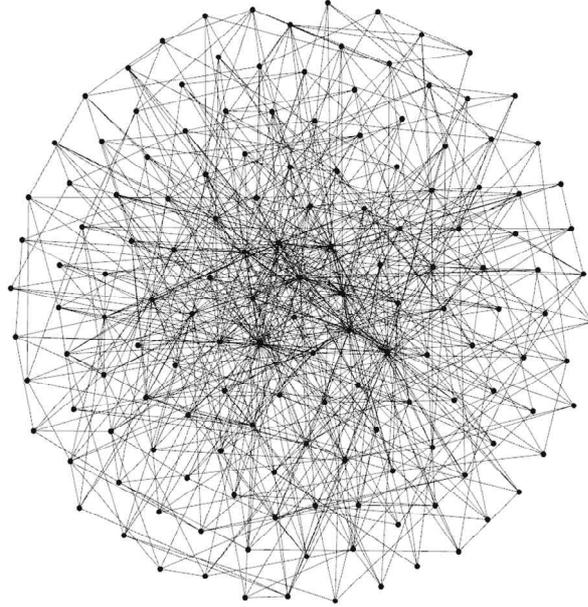
## 2.3 감마 분포의 형상계수 시뮬레이션

무작위 네트워크는 정규분포를 가지고 허브 네트워크는 멱급수 분포를 가지지만 실제 세계의 다양한 문제들은 두 가지 분포를 정확히 따르지 않고 두 분포가 적절히 혼합되어 있다. 두 가지의 서로 다른 분포를 하나의 분포로 설명할 수 있다면 실세계의 분포에 대한 분석이 가능하다.

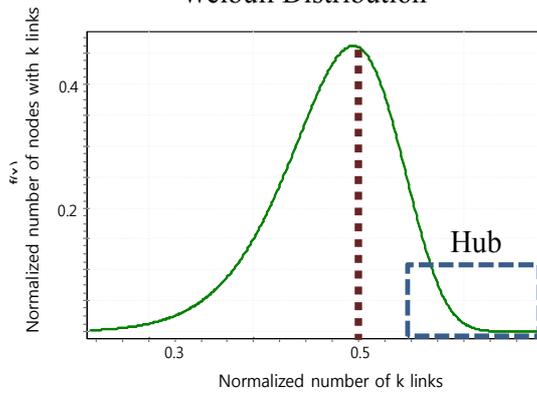
다양한 분포 중 지수분포로부터 파생된 와이블 분포와 감마분포가 형상계수(Shape Factor)  $k$ 와 규모계수(Scale Factor)  $\theta$ 에 따라서 멱급수 분포와 정규 분포의 형태를 동시에 나타낼 수 있다. 150개 절점과 745개의 연결선을 가지는 혼합 네트워크를 각 분포로 각각 시뮬레이션을 해보면 Fig. 2.3과 같다.

혼합 네트워크는 연결선을 독점하고 있는 허브 절점을 가지는 허브 네트워크와 대부분의 절점이 동일한 연결선 수를 가지는 균등한 무작위 네트워크가 혼합되어 있다. 절점이 가지는 연결선수는 두 분포 모두 0에서 1의 범위로 정규화 하였다. 최빈값을 기준으로 시뮬레이션하기 때문에 x축 0.5를 기준으로 두 분포가 유사한 형태를 보이지만 허브 절점을 나타내는 영역에서는 다른 형태를 보인다.

Real-world connection



Weibull Distribution



Gamma Distribution

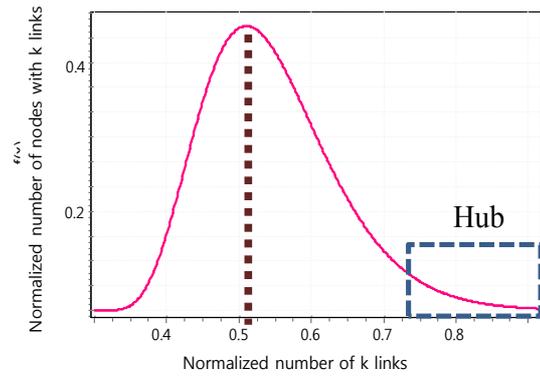


Fig. 2.3 Comparison with Weibull distribution and Gamma distribution

와이블 분포는 최빈값 이후의 영역에서 곡선이 급하게 감소하기 때문에 허브 절점이 분포해 있는 영역을 표현하는데 한계가 있는 반면에 감

마 분포는 완만한 감소 곡선이기 때문에 허브 절점이 분포되어 있는 영역에 대한 정밀한 표현이 가능하다.

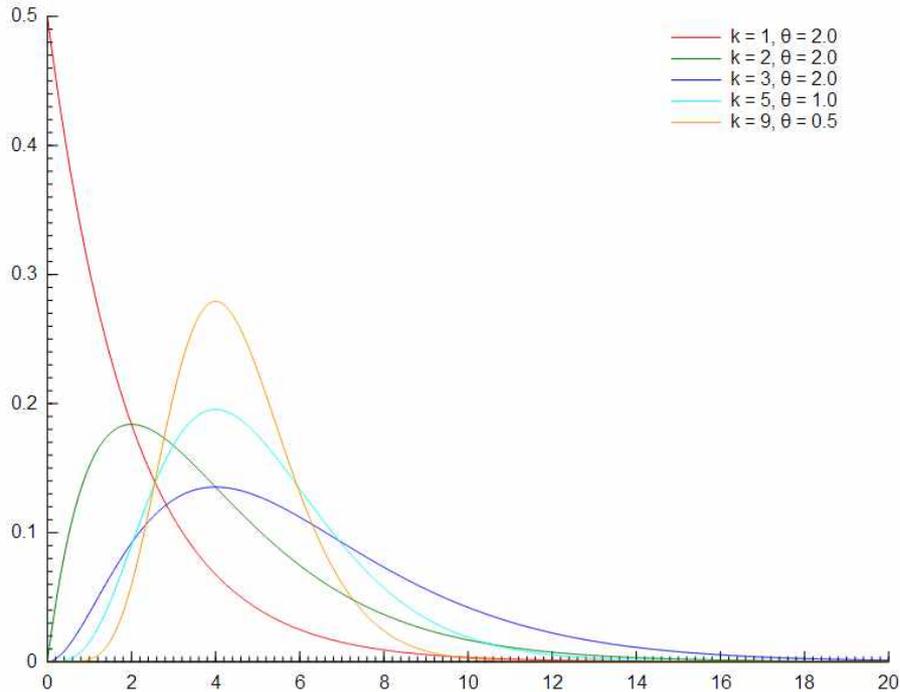


Fig. 2.4 Gamma distribution with scale and shape factor

따라서 와이블 분포가 아닌 감마분포를 이용해서 허브를 가진 혼합 네트워크 및 허브 네트워크를 모의할 때 정확성이 더 높다. 그리고 감마분포는 규모계수와 형상계수만을 이용해서 분포의 평균값, 최빈값, 연결선 수의 최소, 최대범위를 나타낼 수 있기 때문에 분포를 분석할 때도 보다 효율적으로 활용할 수 있다.

규모계수가 평균값으로의 분포에 대한 집중도를 설명한다면 형상계수는 멱급수와 정규분포의 형태를 결정해주는 척도이다. 임의의 네트워크 연결선 수 분포를 감마분포로 모의하면 형상계수와 규모계수가 그 네트워크의 특징적인 고유값이 된다. 고유값을 이용하면 무작위 네트워크와 허브 네트워크의 연결선 수 분포를 비교할 수 있으며 두 네트워크가 혼

합되어 있는 네트워크도 해석할 수 있다.

### 2.3.1 무작위 네트워크

무작위 네트워크(Random network)의 절점 100개를 임의로 생성했고 각 절점들이 다른 99개의 절점들과 연결선을 가질 확률을 0.05로 가정하였다. Fig 2.5 (a)의 무작위 네트워크의 총 연결선은 500개이다. 절점과 연결선수의 분포는 가우시안 분포를 가지며 평균적으로 4~5개의 연결선을 가진 절점이 가장 많이 존재한다.

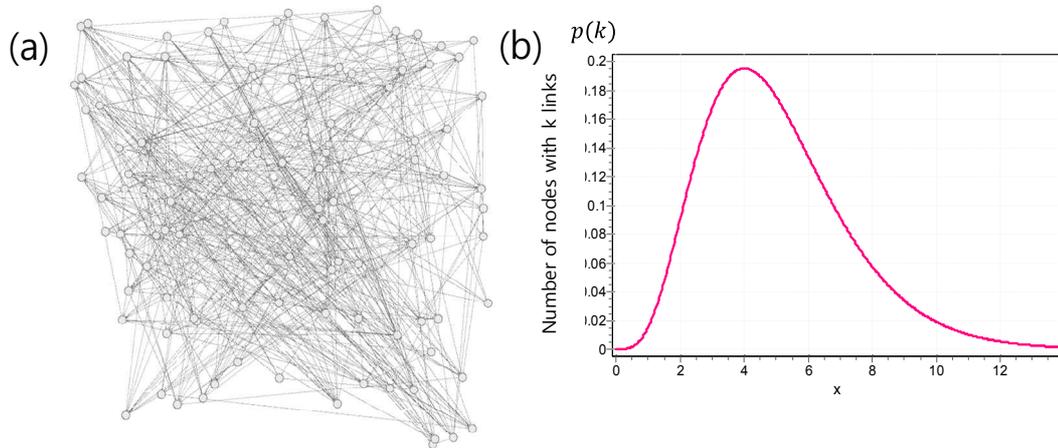


Fig. 2.5 Random network and connectivity distribution

절점 수의 분포를 감마분포로 모의하면 Fig 2.5의 (b)와 같은 가우시안 분포가 나오고 형상계수  $k$ 는 4.17이고 규모계수  $\theta$ 는 1.20이다. Fig 2.5의 (a) 네트워크는 이 두 형상계수와 규모계수의 값을 고유값으로 가진다.

### 2.3.2 허브 네트워크

허브 네트워크(Hub network)는 절점 100개를 Barabasi의 선호적 연결(preferential attachment) 방법으로 자연 발생하며 생성하였다 (Barabási and Albert, 1999). Fig. 2.6 (a)의 허브 네트워크는 Fig. 2.5 (a)의 무작위 네트워크와 같이 연결선이 500개이다. 하지만 무작위 네트워크와 달리 100개의 절점 가운데 최대 38개의 연결선을 독점하고 있는 절점이 존재하고 대다수의 절점은 1~2개의 절점을 가지는 멱급수의 법칙을 따른다.

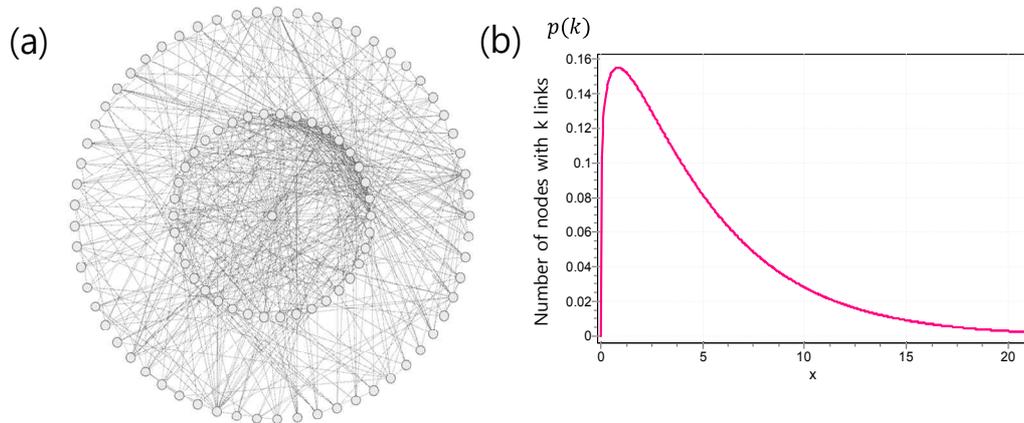


Fig. 2.6 Scale-free network and connectivity distribution

Fig 2.6 (b)는 허브 네트워크의 연결선 수 분포를 감마분포로 모의한 것으로 형상계수  $k$ 는 1.24이고 규모계수  $\theta$ 는 4.03이다. Fig 2.5의 (b)의 가우시안 분포와 다르게 멱급수의 분포를 보이며 형상계수의 값을 비교하면 약 1/3 정도의 크기를 가지고 규모계수는 3배 정도 크다. Fig 2.6 (a)의 허브 네트워크는 Fig 2.5의 (a)의 무작위 네트워크와는 다른 고유값을 가진다.

### 2.3.3 혼합 네트워크

무작위 네트워크와 허브 네트워크가 혼재되어 있는 혼합 네트워크 (Real-world network)는 실제 세계에서 가장 많이 존재하는 분포이다. Fig. 2.5 (a), Fig. 2.6 (a)와 같이 100개의 절점으로 생성하되 그 중 50개는 Fig. 2.5 (a)와 같이 0.05의 확률로 연결선을 생성했고 나머지 50개의 절점들은 Fig. 2.6 (a)와 같이 선호적 연결을 이용하여 네트워크를 구성하였다. Fig 2.7 (a)의 혼합 네트워크는 연결선의 수가 500개 이고 3~4개의 가지는 연결선을 가지는 절점이 가장 많지만 무작위 네트워크와는 다르게 24개 연결선을 가지는 허브 절점도 동시에 존재한다.

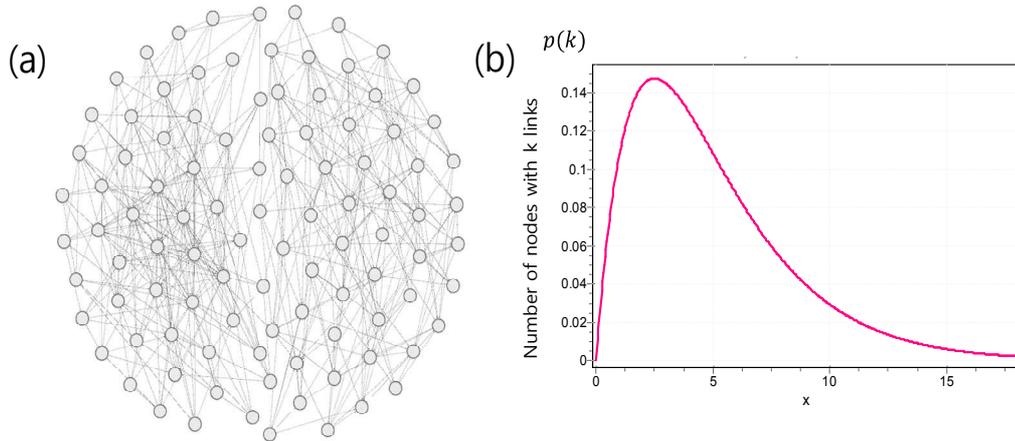


Fig. 2.7 Real-world network and connectivity distribution

Fig 2.7 (b)와 같이 감마분포로 모의했을 때 정규분포의 축과 멱급수 분포의 축 사이에 축을 가지는 분포가 나타난다. 이 분포의 고유값, 형상계수  $k$ 는 1.87이고 규모계수  $\theta$ 는 2.67이다. 혼합네트워크의 형상계수와 규모계수는 각각 허브 네트워크보다는 크고 무작위 네트

워크보다는 작은 값으로 두 네트워크의 형상계수와 규모계수 값의 범위 사이에 있다.

서로 다른 특성을 가진 각 네트워크는 연결선수 분포에 따라서 고유의 형상계수와 규모계수의 값을 가진다. 규모계수는 절점이 가지는 연결선수의 범위를 의미하기 때문에 허브 절점이 가진 연결선수에 대한 정보를 나타내고 대다수의 절점들에 대한 정보, 최빈값은 형상계수를 통해서 나타난다. 동일한 절점과 연결선수를 가진 네트워크에서 두 계수의 곱은 절점이 가지는 연결선의 평균으로 일정하기 때문에 둘의 관계를 이용하면 혼합 네트워크에서 허브 네트워크를 추출할 수 있는 기준을 정립할 수 있다.

## 2.4 혼합 네트워크의 허브 네트워크 추출

### 2.4.1 네트워크의 이산 계층화

네트워크에서 절점은 개체를 나타내며 연결선은 절점간의 관계를 의미한다. 서로 다른 특성을 가진 절점이 혼재되어 있거나 다른 매개체를 이용한 관계가 동일한 연결선으로 정의되어 있다면 다양한 네트워크 해석을 수행할 수 있을지는 몰라도 특성이 다른 점들을 대상으로 중요도를 평가하거나 군집화 해석을 하는데 오류가 발생할 수밖에 없다. 따라서 동일한 특성을 가진 절점들을 추출하는 전처리 작업이 수행되면 네트워크 해석의 효율성 및 해의 정확성이 향상될 수 있다.

Fig. 2.8의 (a)는 연결성이 동일한 15개의 절점과 다른 절점에 비해 연결성이 큰 한 개의 절점 등 총 16개의 절점들이 복잡하게 얽혀있는 임의의 혼합 네트워크이다. 혼합 네트워크에서 연결성이 큰 허브 절점을 탐

색하기 위해서는 16개의 모든 절점들의 연결선을 분석해야 한다. 그러나 (b)와 같이 (a)의 허브 네트워크를 추출하면 15개 절점들이 균등하게 연결되어 있던 무작위 네트워크가 모두 제거되기 때문에 네트워크 해석의 효율성이 증대된다.

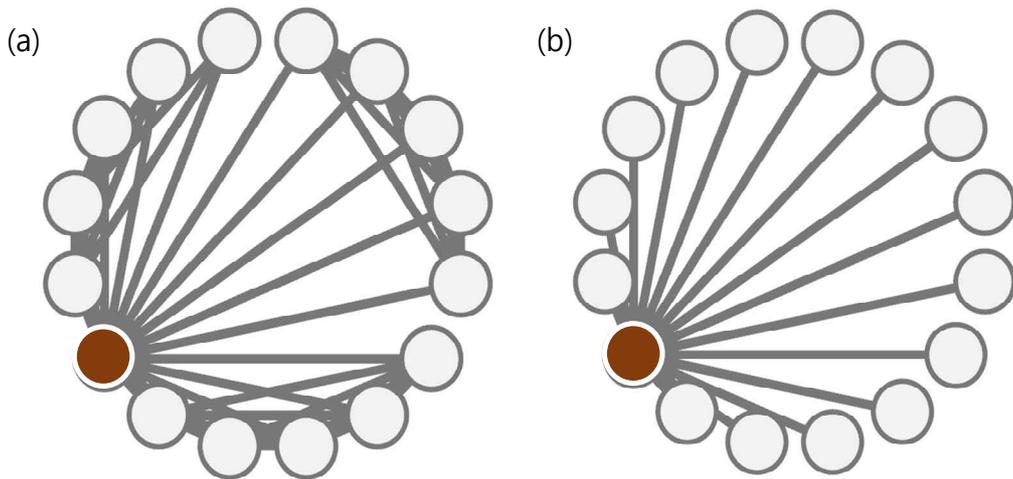


Fig. 2.8 Extraction of hub network in real world connection

동일한 특성을 가진 절점과 연결선으로 구성된 네트워크에서도 연결선 값의 차이가 있다면 그 값에 따라서 계층을 구분할 수 있고 세밀한 분석을 수행할 수 있다. 연결선의 값이 큰 네트워크를 추출하여 거시적인 분석을 수행하고 값이 작은 네트워크는 미시적인 분석을 이용하는 등 분석 방법도 다를 수 있다. 연결성이 높은 절점이 서로 연결된 네트워크는 전체 네트워크의 주요 흐름을 지배하고 있는 허브 네트워크이며 전체 구조를 대표한다. 반면에 연결성이 낮은 절점들로 구성된 무작위 네트워크는 전체에 미치는 파급력은 작지만 주변 절점들과 균집화 될 수 있다. 이산화통한 계층 구분이 이루어지면 서로 다른 특성을 가진 네트워크로 각각 추출할 수 있다.

혼합 네트워크의 계층을 구분하기 위해서는 먼저 절점과 연결선의 특성에 따라서 구분하는 것이 우선이고 동일한 특성을 가진 단일 네트워크 (Homogeneous network)라면 절점과 연결선의 중요도 차이에 따라서 구분할 수 있다. 절점의 중요도는 일반적으로 절점이 가진 연결선에 따라서 결정되기 때문에 연결선의 위상을 우선적으로 분석하는 것이 합리적이다.

#### 2.4.2 혼합 네트워크 임계치

혼합 네트워크에서 서로 다른 특성을 가진 두 네트워크를 추출할 수 있다면 각 네트워크 특성에 맞는 해석을 할 수 있기 때문에 활용도 높은 문제의 해석이 가능하다. 혼합 네트워크의 연결선 수와 추출되는 무작위 네트워크와 허브 네트워크의 연결선 수 총합은 같고 각 네트워크에 존재하는 절점 수는 일정하다. 그리고 연결선 수 분포에서 평균값은 각 절점이 가지는 평균 연결선 수이기 때문에 혼합 네트워크의 평균과 추출된 두 개 네트워크의 평균의 합은 식(1)과 같이 동일하다.

$$\mu = \mu_R + \mu_{SF} \quad (1)$$

$\mu$ : average links of nodes in real-world connections

$\mu_R$ : average links of nodes in random network

$\mu_{SF}$ : average links of nodes in scale-free network

감마분포의 평균은 형상계수  $k$ 와 규모계수  $\theta$ 의 곱과 같기 때문에 네트워크 내 절점이 가지는 연결선 수의 평균은 다음과 같다.

$$\mu = k \times \theta \quad (2)$$

규모계수는 네트워크가 가지는 연결선 수의 범위를 나타내는데, 허브 네트워크는 허브 절점을 독점하고 대다수의 절점들이 연결선 수를 1~2 개를 가지고 있기 때문에 혼합 네트워크의 규모계수와 동일한 너비를 가지고 있다고 가정할 수 있다. 따라서 혼합 네트워크의 규모계수와 허브 네트워크의 규모계수는 동일하다.

$$\theta = \theta_{SF} \quad (3)$$

네트워크의 특성은 연결성을 독점하고 있는 일부 허브 절점들을 통해서 나타난다. 본 연구에서 개발한 허브 네트워크 추출 기준은 혼합 네트워크 연결선수 분포의 규모계수가 허브 네트워크 연결선수 분포의 규모계수와 동일하다고 가정을 하여 산정한다. 규모계수는 허브 절점들의 범위를 나타내는데 그 규모계수가 동일하기 때문에 허브 네트워크가 혼합 네트워크의 허브 절점들을 독점하게 된다. 따라서 추출된 허브 네트워크를 통해서 혼합 네트워크의 특성을 유추할 수 있다.

따라서 형상계수에 따라 두 네트워크의 특성이 결정되는데 평균이 같기 때문에 식(1)에 식(2)와 식(3)을 대입하면 다음과 같다.

$$\begin{aligned} k\theta &= (k\theta)_R + (k\theta)_{SF} \\ k\theta &= k_R\theta_R + k_{SF}\theta \end{aligned} \quad (4)$$

그리고 대다수의 절점이 가지는 연결선 수, 최빈값(mode)은 감마분포에서 식(5)와 같이 구한다.

$$mode = (k-1)\theta \quad (5)$$

허브 네트워크는 허브 절점을 중심으로 방사형의 구조를 가지고 있기 때문에 최빈값을 1로 가정하면,

$$\begin{aligned}(k_{SF}-1)\theta &= 1 \\ k_{SF} &= 1 + \frac{1}{\theta}\end{aligned}\tag{6}$$

식(6)과 같고 허브 네트워크의 형상계수를 산정할 수 있다. 감마분포에서 지수분포를 나타내는 임계가 1이기 때문에 그 값을 기준으로 허브 네트워크의 형상계수,  $k_{SF}$ 가 1과 가까워지면 멱급수와 같은 분포를 보일 확률이 높다.

그리고 형상계수와 규모계수를 곱한 평균값은 네트워크가 가지는 총 연결선 수와 관계가 있기 때문에 혼합 네트워크가 가지는 총 연결선 수는 허브 네트워크와 무작위 네트워크의 연결선 수를 합한 값과 같다. 총 연결선 수를  $link$ 라고 하고 허브 네트워크의 연결선 수를  $link_{SF}$ 라고 하면 연결선 수는 평균에 노드 수  $n$ 을 곱한 값으로 식(7)과 같다.

$$link = k\theta \times n\tag{7}$$

허브 네트워크를 추출할 수 있는 기준이 되는 임계치  $CP_{link}$ 는 연결선 수의 비를 이용하면 식 (8)과 같이 구할 수 있다.

$$CP_{link} = link_{SF}/link\tag{8}$$

식(6)과 식(7)을 식(8)에 대입하면,

$$\begin{aligned}CP_{link} &= \frac{k_{sf}\theta n}{k\theta n} \\ &= \frac{\theta + 1}{k\theta}\end{aligned}\tag{9}$$

식(9)와 같다. 임계치가 정해지면 연결선의 값에 따라서 네트워크를 구분할 수 있고 이를 이용해서 혼합 네트워크에서 허브 네트워크와 무작위 네트워크를 각각 추출할 수 있다.

## 2.5 이산 계층화의 검증

Fig. 2.9 (a)는 간선과 지선으로 각 지역이 연결되어 있는 교통 네트워크를 모델링한 것이다. 간선은 각 지역의 거점을 연결하고 지선은 간선을 기점으로 소지역을 연결하고 있다. 절점은 거점이며 연결선은 간선, 지선과 같은 도로이다. 이 네트워크는 간선과 지선이 혼재되어 있는데 두 연결선의 속도가 상이하기 때문에 매개체가 다른 두 네트워크로 구분할 수 있다. 절점 수가 298개이며 총 2,148개의 연결선을 가진다.

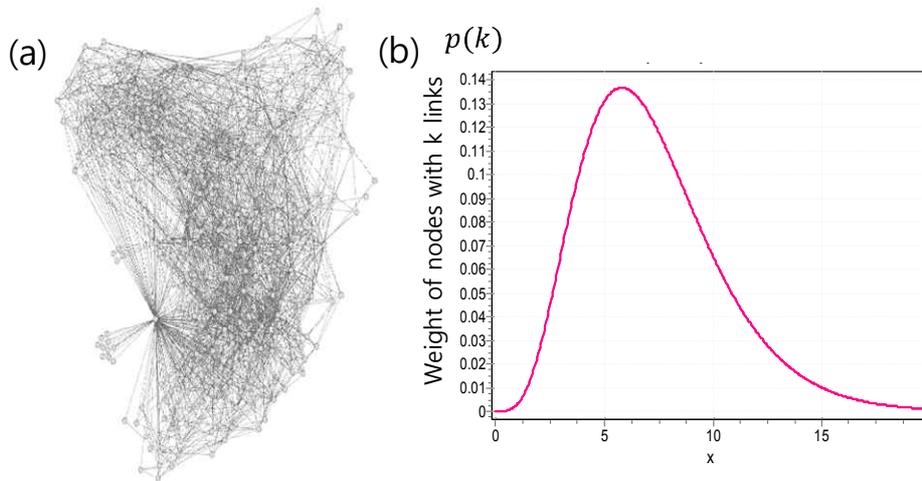


Fig. 2.9 Traffic network and connectivity distribution

네트워크의 연결선 분포를 감마분포로 모의하면 Fig. 2.9 (b)와 같이 형상계수가 5.12, 규모계수가 1.41의 분포로 나타난다. 연결선 수 분포를 보면 대다수의 절점들이 4~6개의 절점을 가지고 있지만 연결선이 집중되어 있는 일부 절점들도 포함되어 있는 것을 확인할 수 있다.

2,148개의 연결선 가운데 504개 연결선의 속도가 다른 간선이기 때문에 식(9)에 따라서 임계치  $CP$ 를 0.334로 하면 298개의 절점과 717개의 연결선을 가진 Fig. 2.10 (a)와 같은 허브 네트워크가 구축된다. 간선 네트워크는 각 지역 거점에 연결선이 집중되어 있는 형태로 구성된다.

이 네트워크의 연결선 수 분포를 감마분포로 모의했을 때 형상계수는 1.778이고 임계치를 이용하여 계산하면 규모계수는 1.41로 동일하고 형상계수는 1.710이다. 실제 분포와 임계 기준으로 추정했을 때를 비교해보면 형상계수 값의 차이는 약 0.068이며 3.81%의 오차가 발생한다. 형상계수는 예측값보다 크게 나왔으며 규모계수는 낮게 산출된다. Fig. 2.10 (a)의 허브 네트워크는 연결선이 불균등하게 집중되어 있는 일부 허브 절점들을 중심으로 형성되어 있으며 20개 이상의 연결선을 가지고 있는 절점도 관찰된다. 이 절점들은 허브 절점들로 다른 절점들에게 큰 영향을 미친다.

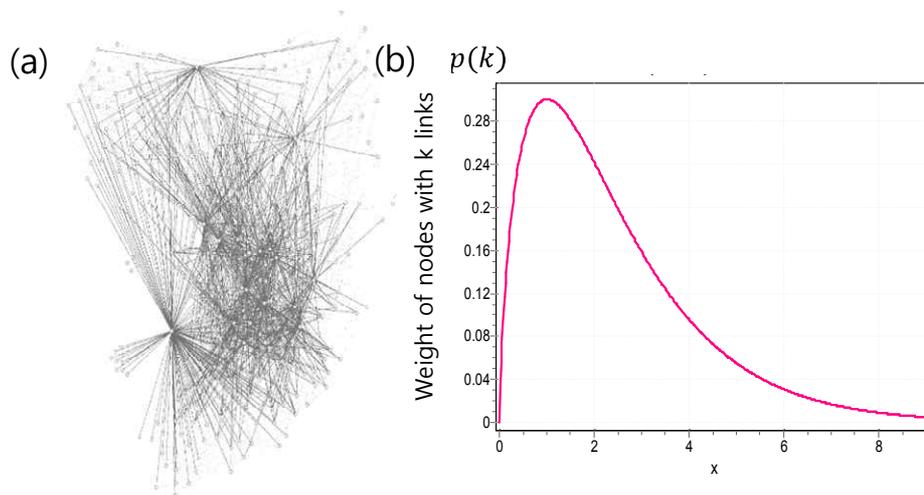


Fig. 2.10 Main line network and connectivity distribution

간선을 제외한 연결선, 1,431개를 이용하여 지선 네트워크를 구축하

면 Fig. 2.11 (a)와 같다. 무작위 네트워크의 연결선 수 분포는 3~4개의 연결선 수를 가진 절점들이 집중되어 있는 정규 분포의 형태를 나타낸다. Fig. 2.9 (a)의 다수의 절점들이 허브 절점들과 연결된 1개의 연결선을 가지고 있기 때문에 이 수를 제외하면 최빈값을 유추할 수 있다.

감마분포의 형상계수를 이용한 임계치 분석의 정확도는 96.19% 이다. 혼합 네트워크의 특성인 연결선 수 분포를 감마분포로 모의하고 그 형상계수를 이용해서 두 네트워크를 추출했을 때 정확도는 실제 분포와 비교할 때 95%이상으로 신뢰할만한 수준이라고 판단된다.

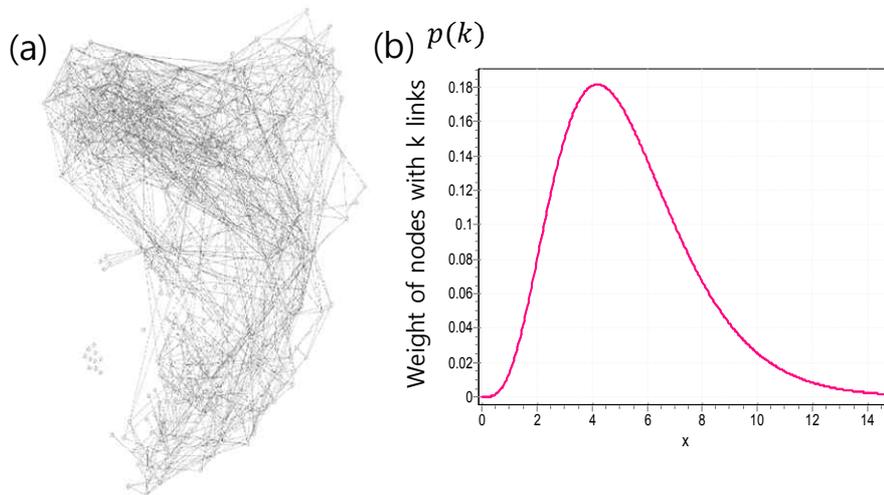


Fig. 2.11 Branch line network and link distribution

이 결과는 연결선의 속도가 다른 제한된 네트워크를 대상으로 수행되어 도출된 결과이다. 허브 네트워크의 추출에서 최빈값을 1로 가정하지 않거나 규모계수를 혼합 네트워크의 값과 동일하게 가정하지 않고 분석적인 방법을 이용해서 적절한 범위를 결정할 수 있다면 혼합 네트워크에서 보다 정확도 높은 각 네트워크의 추출이 가능할 것으로 기대된다.

그리고 무작위 네트워크는 혼합 네트워크의 추출 이후에 결정된다. 혼합 네트워크에서 무작위 네트워크의 형상계수와 규모계수를 추정할 수 있는 근거가 있다면 정확도가 높은 결과를 기대할 수 있을 것이다. 추후 다양한 사례에 적용해서 지속적으로 비교 검증해 나가면 결과의 신뢰도를 보다 향상시킬 수 있을 것으로 판단된다.

## 2.6 소 결

네트워크는 절점들의 연결선 수 분포에 따라서 서로 다른 특성을 가진다. 허브 네트워크는 기하학적인 관계를 모사할 때 이용되는 무작위 네트워크와 달리 주로 위상적인 관계를 모델링하기 때문에 연결선 수 분포가 불균등하며 허브 절점이 존재한다. 허브 절점은 네트워크의 특성을 대표하기 때문에 허브 절점으로 구성된 네트워크는 복잡계 연구에 있어 핵심이 되고 있다. 특성이 다른 네트워크가 혼합된 실제 문제에서 허브 네트워크를 추출할 수 있다면 해석의 효율성과 해의 정확성을 향상시킬 수 있다.

본 연구에서는 허브 네트워크와 무작위 네트워크의 연결선 수 분포를 분석하고 감마분포로 나타내어 형상계수, 규모계수라는 두 모수로 대체하였다. 두 모수를 이용하여 혼합 네트워크에서 허브 네트워크 추출의 임계치를 정립하였으며 이를 교통망 네트워크와 같이 간선과 지선에 따라서 속도가 구분되어 있는 실제 사례에 적용하여 검증하였다. 실제 간선만으로 이루어진 네트워크의 연결선 수 분포를 감마분포로 모의했을 때의 형상계수와 임계치를 이용하여 산정한 형상계수를 비교했을 때 95%이상의 정확도를 가지는 것으로 나타났다.

본 연구는 차원이 다른 두 개의 네트워크가 혼재되어 있을 때 연결선 수 분포라는 특성을 이용해서 허브 절점을 가지고 있는 허브 네트워크와 무작

위 네트워크를 각각 추출하는 방법을 정립하였다. 연결선의 속도가 다른 네트워크에 연결선의 값에 따라서 서열화할 수 있는 기준을 개발한 방법으로 제한적일 수 있지만 허브 네트워크의 형상계수와 규모계수를 추정하는 과정이 보다 분석적인 방법이 적용될 경우 정확도는 보다 향상될 것이며 본격적인 네트워크 분석을 수행하기 전 처리 과정으로 활용도가 높을 것으로 기대된다.

## 참고문헌

1. Albert, R., 2005. Scale-free networks in cell biology. *Journal of cell science* 118(21): 4947-4957.
2. Barabási, A.L. and R. Albert, 1999. Emergence of scaling in Random networks. *science* 286(5439): 509-512.
3. Barabási, A.L., R. Albert and H. Jeong, 2000. Scale-free characteristics of random networks: the topology of the world-wide web. *Physica A: Statistical Mechanics and its Applications* 281(1-4): 69-77.
4. Barthélemy, M., A. Barrat, R. Pastor-Satorras and A. Vespignani, 2004. Velocity and Hierarchical Spread of Epidemic Outbreaks in Scale-Free Networks. *Physical Review Letters* 92(17): 178701.
5. Boccaletti, S., V. Latora, Y. Moreno and M. Chavez, 2006. Complex networks: Structure and dynamics. *Physics reports* 424(4-5): 175-308.
6. Boccara, N and K. Cheong, 1993. Critical behaviour of a probabilistic automata network SIS model for the spread of an infectious disease in a population of moving individuals. *Journal of Physics A: Mathematical and General*, 26(15).
7. Boguná, M. R. Pastor-Satorras and A. Vespignani, 2003. Absence of epidemic threshold in scale-free networks with degree correlations. *Physical review letters* 90(2): 028701.
8. Eguiluz, V.M. and K. Klemm, 2002. Epidemic threshold in structured scale-free networks. *Physical Review Letters* 89(10): 108701.
9. Holger, E., L. Mielsch and S. Bornholdt, 2002. Scale-free topology

- of e-mail networks. arXiv preprint cond-mat/0201476.
10. Khan, H., R.N. Mohapatra, K. Vajravelu and S.J. Liao, 2009. The explicit series solution of SIR and SIS epidemic models. *Applied Mathematics and Computation* 215(2): 653–669.
  11. Moreno, Y., J. B. Gómez and A. F. Pacheco, 2003. Epidemic incidence in correlated complex networks. *Physical review E* 68(3): 035103.
  12. Newman, M. E. J., 2003. The Structure and Function of Complex Networks. *SIAM Rev* 45(2): 167 - 256.
  13. Romualdo, P.S. and V. Alessandro, 2001. Epidemic spreading in scale-free networks. *Physical review letters* 86(14): 3200–3203.
  14. Sema, N., L. Matthews, P.R. Bessell, S. WJ Reid and R.R. Kao, 2011. Generating social network data using partially described networks: an example informing avian influenza control in the British poultry industry. *BMC Veterinary Research*, 7(66).
  15. Sterman, J.D., 1994. Learning in and about complex systems. *System Dynamics Review* 10(2–3): 291–330.
  16. Wang, X.F. and G. Chen, 2003. Complex networks: small-world, scale-free and beyond. *Circuits and Systems Magazine IEEE*, 3(1): 6–20.



## 제 3 장 허브 네트워크의 취약성 분석

### 3.1 서론

무작위 네트워크와 다르게 허브 절점을 가지고 있는 허브 네트워크는 웹 구조, 생물 구조, 사회적 관계 구조 등에서 쉽게 발견되며 각각의 사례가 가진 문제에 대해서 다양한 시각으로 해석되고 있다 (Aldana and Cluzel, 2003; Huberman, 2001; Fell and Wagner, 2000; Albert et al., 1999; Bearman et al., 2004). 일부 허브 절점들이 네트워크의 대다수 연결선을 지배하고 있는 불균등한 구조 때문에 자연스럽게 절점들은 서로 다른 중요도를 가진다. 허브 절점들은 다른 절점들보다 연결선을 많이 가지고 있기 때문에 네트워크 구조 전체에서 중요한 역할을 한다. 절점과 연결선이 지속적으로 증가하는 성장 네트워크에서 생성되는 절점들은 허브 절점들과 이어진 연결선을 먼저 가질 확률이 가장 크다. 따라서 시간이 지나면서 절점들의 규모는 점차 커지고 허브 네트워크의 특성이 더욱 강화된다.

허브 절점은 네트워크에서 중추적인 위치에 있기 때문에 외부에서 똑같은 공격을 받게 되더라도 다른 절점과 비교할 때 전체 구조에 미치는 영향력이 크다. 다행스럽게도 허브 절점들은 전체 절점 가운데 아주 극소수이기 때문에 전체 절점들이 공격을 받을 확률이 모두 동일하다면 허브 네트워크는 쉽게 붕괴되지 않는 견고한 구조를 가진다 (Albert et al., 2000; Tu, 2000). 장애에 대한 저항력이 높아야만 하는 시스템은 선천적으로 불균등한 연결선 수 분포를 가진다. 생물학적 종은 세포들이 그물처럼 얽혀있지만 중요도가 높은 세포들을 중심으로 불균등하게 연결되어 있기 때문에 견고한 구조로 구성되어 있다 (Dunne et al., 2002; Sole and Montoya, 2001; Jeong et al., 2001).

그러나 허브 절점들이 막대한 영향력을 가지기 때문에 집중적으로 공격을 받는다고 하면 허브 네트워크는 가장 위험한 구조가 되기도 한다. 소수의 허브 절점들로 전체 연결성이 유지되고 있는 구조에서 허브 절점들이 파괴되면 대다수의 절점들의 연결성이 끊어지고 의미 없는 구조가 된다. 인접한 절점들 간의 다양한 연결성은 전체 구조에서 여분성(redundancy)을 가진다 (Cohen et al., 2000; Callaway et al., 2000). 연결성이 크지는 않지만 소수의 연결선을 가지기 때문에 네트워크는 일부 절점들이 제거되더라도 대체할 수 있는 다른 경로를 통해서 붕괴되지 않는다. 중추적인 허브 절점들이 지배하는 치명적인 경로가 아니라면 대부분 공격에는 대처할 수 있는 대안적인 경로를 가지고 네트워크의 장애나 공격에 대응한다. 모든 절점이 같은 확률을 가지는 내부에서 발생할 수 있는 장애에 대해서는 관리성이 뛰어나지만 허브 절점에 대한 외부의 공격에는 취약하기 때문에 네트워크에서는 견고성과 취약성이 공존한다 (Albert et al., 2000; Aldana and Cluzel, 2003).

바이러스나 질병의 확산과 관련된 문제에서도 네트워크의 취약성과 견고성에 대한 연구가 있다 (Klov Dahl et al., 1994; White et al., 1986; Gupta et al., 1989). 확산의 대상은 절점, 확산 경로가 될 수 있는 대상 간의 연계요인을 연결선으로 추상화시켜 네트워크 분석을 통해 확산 경로를 추적하거나 차단한다. 네트워크 분석은 확산 해석의 적합한 도구로 이용될 수 있다. 특히 허브 네트워크 구조에서 발병하면 대다수의 절점들에게 영향력을 미치는 허브 절점들이 존재하기 때문에 임의의 절점들에게 대책을 집중하는 것보다 허브 절점들에게 바이러스의 차단을 위한 대책을 집중하는 것이 전체 구조의 붕괴를 막을 확률이 높다 (Albert et al., 2000; Cohen et al., 2001). 따라서 질병과 질병의 매개체로 구축된 네트워크에서 허브 절점을 구하게 되면 확산 방지 및 예방 대책을 수립하기 위한 기초자료로 활용될 수 있다.

본 연구에서는 허브 네트워크에 존재하는 허브 절점의 탐색 방법에 대

해서 고찰하고 그 방법을 가축 전염병 가운데 조류 인플루엔자(Highly Pathogenic Avian Influenza, HPAI) 확산 사례에 적용하고자 한다. HPAI와 관련된 자료를 이용해서 네트워크를 구성하고 허브 절점 탐색 방법을 적용해서 확산 차단을 위한 가장 효율적인 대책을 제시하고자 한다.

### 3.2 취약성 분석

허브 네트워크는 연결선 수 분포가 식(1)과 같은 멱급수의 법칙을 가진다.

$$p(k) \sim k^{-\gamma} \quad (1)$$

$p(k)$  : probability that a node has  $k$  links

$k$  : number of links

$\gamma$  : degree exponent

연결선의 분포가 완만하게 감소하는 멱급수의 그래프를 따르기 때문에 특징적으로 연결선을 많이 가지는 허브 절점이 존재한다는 것을 보여준다. 허브 네트워크는 다른 수많은 절점과 비교할 때 상대적으로 연결선 수를 많이 가지는 소수의 허브 절점들을 가지게 되고 불균등한 분포를 나타낸다.

네트워크에서 허브 절점은 구조적 안정성, 견고성, 취약성 등을 해석하는데 이용된다. 허브 네트워크의 절점들은 연결선 수에 따라 중요도를 가진다. 절점과 연결선의 조합으로 이루어진 네트워크에서 두 가지 요인만을 가지고 중요도를 정량적으로 산정할 수 있는 방법을 중심성 분석(Centrality Analysis)이라고 한다. 네트워크에서 중심성 분석은 핵심적 위치에 있는 절점들의 특성을 값으로 표현하여 준다. 그 중 연결 중심성

(Degree Centrality)은 1개의 절점과 연결되어있는 연결선의 수를 계산하며 연결선 수 분포에서도 가장 많이 활용되고 쉽게 산정할 수 있는 방법이다 (Beauchamp, 1965). 연결 중심성을 이용하면 다른 요소들과 연결성이 좋은 허브 역할을 하는 절점들을 찾아낼 수 있다.

$$(\text{Degree Centrality})_i = \frac{k(n_i)}{g-1} \quad (2)$$

식 (2)에서  $g$ 는 네트워크 내 절점간의 총 연결선 수이며  $k(n_i)$ 는  $i$ 번째 절점의 연결선 수를 의미한다. 연결선마다 가중치가 있는 가중네트워크에 적용한다면  $k(n_i)$ 은  $i$ 번째 절점의 연결선 수의 값을 모두 합한 값이다.

근접 중심성(Closeness Centrality)은 네트워크 내에서 기하학적으로 가장 중심에 있는 절점을 구하는 방법으로 모든 절점 간의 네트워크 직경과 비교해서 계산한다. 매개 중심성(Betweenness Centrality)은 1개의 절점이 다른 2개의 절점 간의 최단 경로에 놓이게 되는 비율을 합하여 계산하는 방법이다 (Loomis, 1950; Freeman, 1979).

$$(\text{Closeness Centrality})_i = (g-1) \left[ \sum_{j=1}^g d(n_i, n_j) \right]^{-1} \quad (3)$$

식 (3)의  $d(n_i, n_j)$ 는  $i$ 번째 절점과  $j$ 번째 절점간의 기하학적인 거리를 뜻한다. 매개 중심성은 네트워크 내의 다른 두 노드의 연결 경로에 놓여 있는 정도를 측정해서 연결성이 좋은 절점을 선택한다.

$$(\text{Betweenness Centrality})_i = \sum_{j < k}^g g_{jk}(n_i) / [(g-1)(g-2)/2] \quad (4)$$

$$g_{jk}(n_i) = 0, \text{ if nodes } n_i \text{ is not between } n_j \text{ and } n_k ,$$

$$1, \text{ if nodes } n_i \text{ is between } n_j \text{ and } n_k$$

$m$  : number of nodes

식 (4)의  $g$ 는 식 (2), (3)의  $g$ 와 같이 네트워크 내 절점간의 총 연결선

수를 나타내고  $g_{jk}(n_i)$ 은 네트워크 내  $j$ 번째 절점과  $k$ 번째 절점을 최단 거리로 연결했을 때  $i$ 번째 절점의 통과 여부를 측정한다. 통과하면 1, 통과하지 않으면 0의 값을 준다. 매개 중심성은 절점간 최단 경로를 파악해야하기 때문에 경로 추적에서 중추적인 역할을 하는 절점을 찾아내는데 유효하다.

세 가지 방법이 적용되는 네트워크는 양방향성, 일방향성을 가지는 네트워크의 모든 경우에 적용이 되며 일방향성을 가질 경우에는 절점으로 들어오는 연결선과 나가는 연결선을 구분하여 중심성 지수가 산정된다. 허브 네트워크의 중심성을 산정할 때 세 가지 지수는 서로 보완적인 역할을 통해서 네트워크 내의 주요 절점을 선정하고 취약성을 분석하는데 효율적이고 신뢰성 높은 결과를 보여줄 수 있을 것으로 판단된다. 허브 절점들의 역할이 집중되어 있는 네트워크에서 절점들의 중심성 분석은 확산 문제의 해석에 있어서 유용하게 이용될 수 있다. 상대적으로 연결성이 좋은 중심성 지수가 높은 절점에 확산 예방 대책을 집중한다면 효율적으로 확산 속도를 늦추거나 전체 네트워크 구조가 붕괴되는 것을 막을 수 있다.

### 3.3 네트워크 취약성 분석을 이용한 확산 해석

#### 3.3.1 네트워크 허브와 확산

네트워크에서 허브 절점을 이용한 확산 문제 해석은 지속적으로 연구되고 있다 (Moreno et al., 2004; Newman, 2002; Liljeros et al., 2001). 복잡계 네트워크 연구에서는 확산의 대상을 절점, 확산 경로가 될 수 있는 대상간의 연계요인을 연결선으로 추상화시키고 네트워크 분석을 통해 확산 경로를 추적하거나 차단한다. 절점과 절점간의 연결선으로 복잡하

게 얽힌 시스템의 일부에서 확산이 시작되면 시간이 지나면서 영향을 받는 절점이 기하급수적으로 증가하고 확산 속도도 커지는 양상을 보인다. 확산의 주체가 전염병일 경우 예측할 수 없는 확산의 방향과 빠른 속도는 막대한 물적, 인적 피해를 야기한다. 다른 절점들보다 연결성이 좋은 허브 절점은 확산에 있어서도 중추적인 역할을 하기 때문에 허브 절점을 찾아서 우선적으로 방역하면 확산 속도를 늦추거나 차단할 수 있다. 확산 문제에 네트워크 허브 절점 분석은 다양하게 이용될 수 있다 (Liljeros et al., 2001; May and Anderson, 1989; Moreno et al., 2004).

### 3.3.2 모델의 적용

전염성이 강한 질병의 확산은 경제적인 피해뿐만 아니라 인간의 생명에게도 큰 위협을 주고 있을 정도로 심각한 문제가 되고 있다. 사회구성원간의 복잡하게 얽혀있는 관계에 따라서 질병이 발병할 경우 사회 전반으로 급속도로 확산된다. 또한 환경 변화와 함께 바이러스의 변이가 다양화되며 백신을 통한 질병의 적절한 대응에 어려움을 겪고 있다. 전염병과 함께 가축전염병의 유행도 증가하고 있다. 특히 HPAI 바이러스(H5N1)는 1997년 농가와 가축 거래 시장(live animal market)에서 발병된 이후 60개국 이상의 지역에서 반복해서 발병하고 있다. 이는 가금류에게 치명적인 치사율을 가지며 가금산업에 막대한 경제적 손실을 유발하며 최근엔 동물에서 인간으로 전염되기도 한다 (Martinez, 2011, WHO, 2010, WHO, 2012). 그리고 한 국가의 일부 지역에서 발병하더라도 철새를 통해서 전파될 수 있기 때문에 인접 국가에 동시에 위협이 될 수 있다. 한국에서는 2003년 최초 발병 이후 총 219건의 발병 사례가 보고되는 등 HPAI 예방대책이 시급한 문제로 인식되고 있다. 특히 한국에서 일어난 2008년 발병은 생산단계 및 유통단계 피해액을 모두 합산하면 약 6300억(63 million dollar)의 막대한 경제적 피해

를 입은 것으로 추정하고 있다 (Woo et al., 2008). HPAI는 가축을 대상으로 하는 질병이기 때문에 최초 발병시 잠복기 및 축주가 발병을 인지하는데 걸리는 시간 등으로 초기 확산에 대한 추적이 어려우며 다양한 변이 때문에 근본적인 예방 백신을 이용한 확산 차단이 불가능하기 때문에 방역대책에 어려움을 겪고 있다.

현재 가축전염병의 대책은 발병 농가와 발병 농가의 일정 반경 내 농가를 모두 살처분하는 방식이다. 그러나 발병한 농가와 기하학적으로 근접한 모든 농가를 폐쇄함에도 불구하고 현재 바이러스의 추적이 쉽지 않으며 유행할 때 확산을 막기 어렵다. 현재의 방역대책은 바이러스를 전달할 수 있는 농가를 방문하는 매개체에 대해서 경로 추적을 하지 않을 뿐만 아니라 발병 지역의 기상 및 지형적 조건을 고려한 공기를 통한 간접적 확산에 대한 고려 없이 일정 반경 내의 농가를 폐쇄한다. 확산 요인에 대한 정확한 분석 없이 해당 농가의 일정 반경 내 위치한 근린 농가를 완전 폐쇄하는 대책은 비효율적인 방안일 수밖에 없다. 따라서 HPAI의 확산 경로가 될 수 있는 매개체를 분석하고 피해를 최소화할 수 있는 지점에 관리를 집중시켜 초기 확산에 효율적으로 대처하는 방역 대책이 필요하다.

HPAI 확산 해석을 위해 농가를 방문하는 가금관련업체 차량 및 관계자의 출입에 대한 직접 확산과 공기 중의 간접 확산에 대한 자료를 이용해서 네트워크 취약성에 대한 모델을 포괄적으로 검증 및 적용해 보고자 한다. 2008년 한국 김제 지역 발병사례를 기초자료로 농가들의 업체 이용 자료와 공기 역학 자료를 바탕으로 HPAI 확산 네트워크를 구성하고 연결선 수 분포를 통해서 네트워크의 구조를 파악하고 그에 적합한 취약성 분석을 통해서 감염에 취약하고 연쇄 사고를 통해서 타 농가에게 영향을 미칠 가능성이 높은 농가를 선정하였다. 그리고 실제 발병 농가와 비발병 농가의 중심성 지수를 비교하고 발병 경로를 해석하여 본 연구에서 제안한 확산 네트워크를 검증하였다. 이를 통해서 기

존의 방역 체계와 달리 일부 농가의 집중 관리를 통해서 초기 확산을 차단할 수 있는 효율적인 방역 대책을 제시할 수 있을 것으로 기대된다.

### 1) 대상자료

HPAI 질병은 2000년 이후 한국에서 '03/'04년, '06/'07년, '08년, '12년 등 이미 네 번의 유행이 있었다. 이 가운데 '08년 한국 김제에서 최초 발병한 HPAI는 전국적으로 42일간 지속되며 전국적으로 33건의 발병이 확인되었다 (APQA, 2008). 다른 세 가지 유행과 비교해서 '08년 HPAI 유행은 발생일수는 감소하고 감염건수는 증가하는 등 단기간에 확산이 집중되는 양상을 보였다. 그리고 감염건수도 많고 관련 농장에 대한 역학조사 내용도 양적으로 다양하기 때문에 통계적 조사를 수행하는데 있어서 모델의 정확도를 높일 수 있을 뿐만 아니라 검증이 가능한 사례가 될 수 있다.

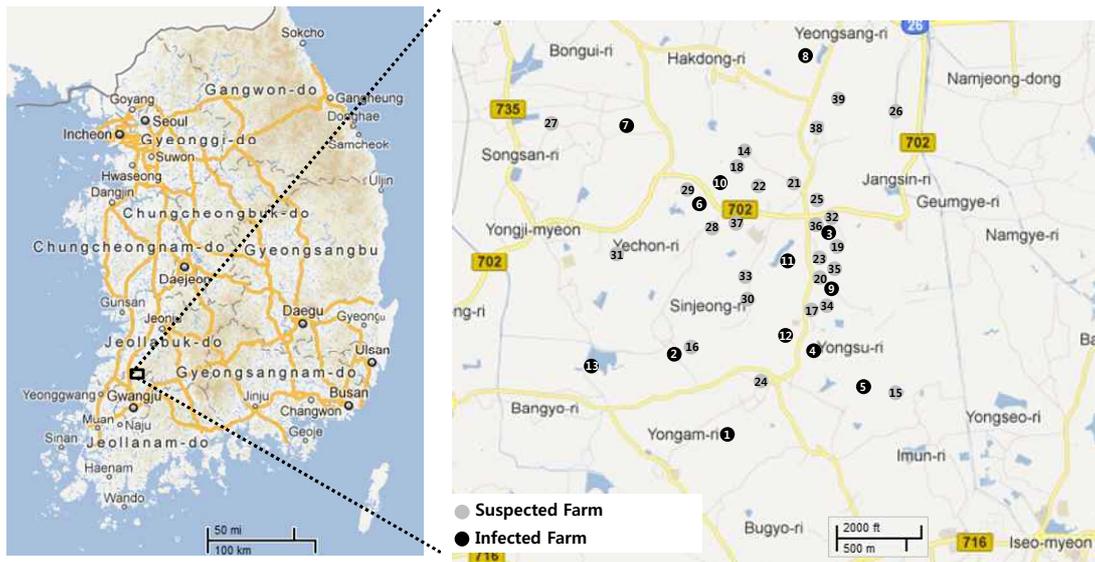


Fig. 3.1 Location of 39 farms in Gimje, South Korea.

최초 발병 지역인 김제 지역은 Fig. 3.1에서 보듯이 총 39개 농가 가운데 12개 농가가 감염되었다. 발병농가는 발병 순서에 따라서 1번에서 12번까지 번호를 줬으며 비발병농가는 13번부터 39번까지 각각 번호를 부여하였다. 이 지역은 주변에 가금 관련 대기업을 위치한 농가 밀집 지역으로 확산시 피해 규모가 커지며 다른 지역으로의 파급력이 큰 HPAI 주요 위험 지역으로 볼 수 있다. 반경 2km이내에 39개 농가가 밀집된 지역으로 공기중 직접 확산의 영향이 있을 뿐만 아니라 동일한 가금관련 업체를 이용하는 농가들이 많고 차량 및 관계자들의 왕래가 빈번한 특징을 보인다. 본 연구에서는 최초 발병 지역인 김제 지역의 39개 농가를 대상으로 HPAI 확산을 분석할 것이다.

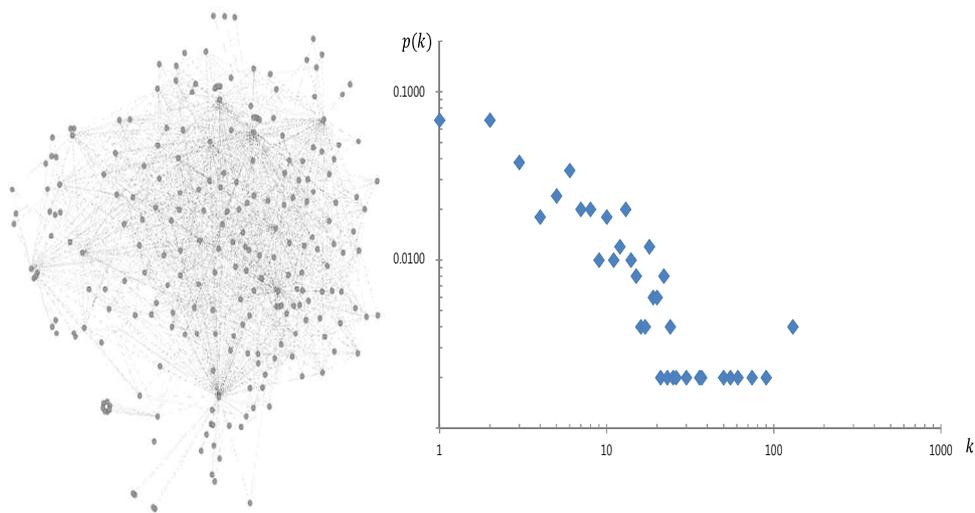


Fig. 3.2 Link distribution on poultry related business network

## 2) 네트워크 구축

2008년 발병 당시 검역기관은 총 전국 11개 시·도 19개 시·군·구의 235개의 가금 농가를 대상으로 방문 검역 조사를 수행하였다. 조사된 자

료 가운데 농가의 업체 이용 자료를 기반으로 네트워크를 구축하면 Fig. 3.2와 같고 연결선 수 분포는 일부 허브 절점들에 연결선이 집중되어 있는 허브 네트워크의 형태로 구축된다. 가금업체 네트워크에서는 일부 농가들이 많은 고객을 가진 업체나 다수의 업체를 동시에 이용함에 따라 절점들의 연결선 수가 불균등해지고 일부 허브 농가들이 생겨난다.

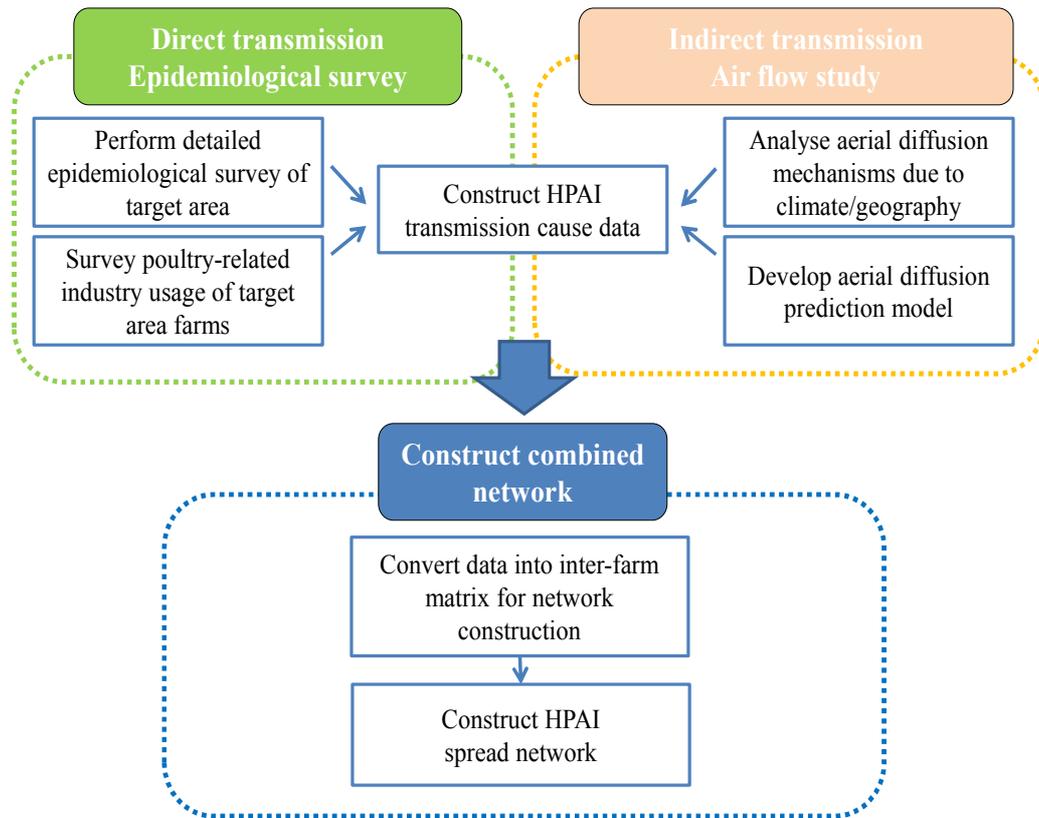


Fig. 3.3 Data collection for consideration of direct spread by poultry-related business vehicles or personnel visiting farms and indirect spread by transmit through air

김제 지역의 자료를 이용하여 Fig 3.1의 39개 농가를 Fig. 3.4의 (a), (b)와 같은 약품 네트워크, 사료 네트워크 등 2개의 직접 HPAI 확산 네트워크로 각각 구성하였다. 대상지역은 가금농가들이 밀집된 지역이기 때문에 공기를 통한 HPAI 바이러스의 간접적 확산이 충분히 가능한 지역이다. 매개체를

이용한 직접적 확산과 함께 간접적으로 바이러스가 확산되어 가는 경우를 분석하기 위해서 대상 지역의 농가간 공기 영향도를 분석하였다. 발병 당시 지역의 풍향 및 대상 지역의 지형적 요인을 고려해서 공기 확산 메커니즘을 분석하고 전산유체역학 시뮬레이션을 이용해서 공기 확산 예측 모델을 개발하였다. 그 자료를 바탕으로 39개의 농가의 서로간의 농가간의 공기 영향도를 분석한 자료를 구축했고 이를 농가간 공기영향 상관관계를 나타내는 행렬 자료로 변환하고 Fig. 3.4의 (c)와 같이 간접 확산 네트워크인 air network를 구성하였다. 역학 조사와 공기 유동에 관한 조사를 통해서 자료를 수집하고 HPAI 직, 간접 확산 요인에 대하여 각각 구분하고 그 자료를 네트워크에 적합한 자료로 변환하여 HPAI 확산 네트워크를 구성하였다.

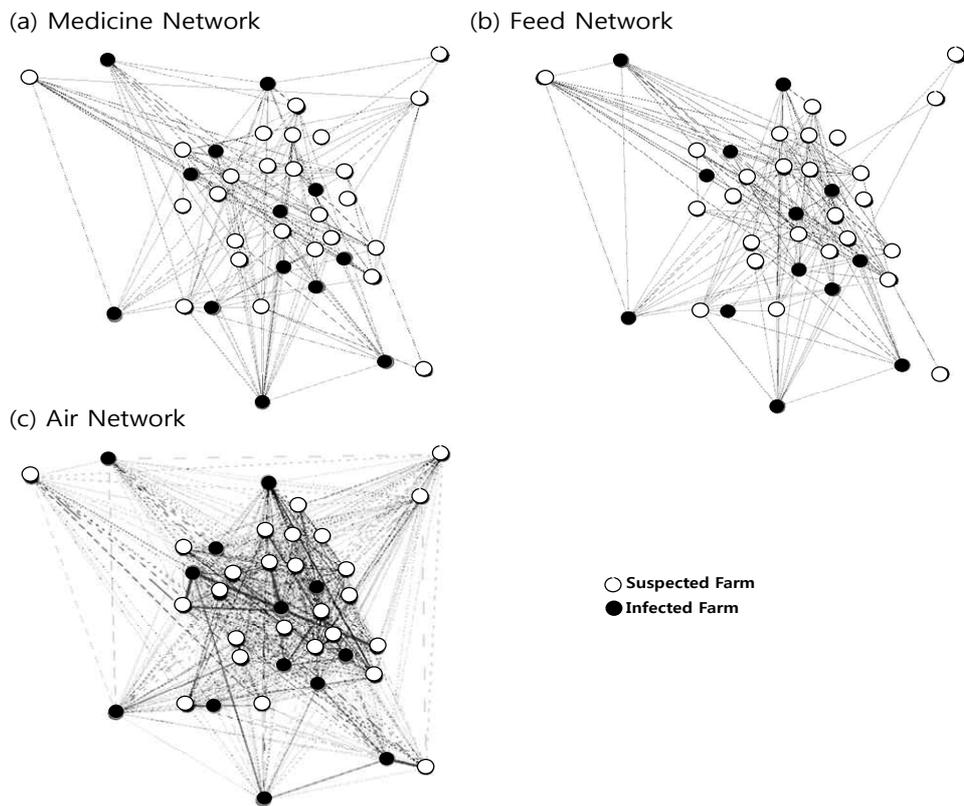


Fig. 3.4 HPAI spread network with poultry related business data;  
 The whole network has 39 nodes, (a) medicine network, (b)  
 direct feed spread network, and (c) indirect air spread network.

직접 확산을 해석할 수 있는 Fig 3.4의 (a) 약품 네트워크와 (b) 사료 네트워크는 동일 업체를 이용하는 농가들끼리 연결되어 있는 완전 네트워크(complete Network)로 구성된다. 두 군데 이상 업체의 관리를 받는 농가들 때문에 완전 네트워크 간의 연결선을 가질 수 있어 39개의 농가는 닫힌 네트워크(closed network)가 된다. 약품 네트워크는 39개의 절점과 288개의 연결선을 가지며 사료 네트워크는 39개의 절점, 284개의 연결선을 가진다. 두 네트워크는 각각 절점당 평균적으로 7.39개, 7.28개의 연결선을 가진다. 각 연결선의 방향성은 양방향성을 가진다.

Fig 3.4의 (c) 공기영향 네트워크는 39개의 절점과 1282개의 연결선으로 구성되어 있으며 절점당 32.87개의 연결선을 가진다. 공기중 바이러스 확산은 한 농가에서 풍향에 따라 다수의 농가에 다발적으로 영향을 미치므로 직접적 확산과 비교해서 상대적으로 연결선수가 많다. 그리고 직접 확산 네트워크의 연결선의 값이 모두 동일한데 반해서 간접 확산 네트워크에서는 농가간 공기 중 영향의 크기가 매우 다양해서 가장 약한 연결선과 가장 강한 연결선은 약 8600만배의 차이가 나는 가중 네트워크로 구성된다.

본 연구의 전개과정을 나타내면 Fig. 3.5와 같다. Fig 3.4의 구축된 네트워크 모델 중 직접 확산 네트워크 모델에 연결 중심성, 근접 중심성, 매개 중심성의 3가지 중심성 분석을 하고 가중 네트워크인 간접 확산 네트워크 모델에 연결 중심성 분석을 수행하였다.

중심성 분석 수행 후 모델의 검증을 위해서 실제 '08년 실제 발병 농가와 비발병 농가의 중심성 값의 비교를 통해서 실제 발병과 네트워크 중심성의 상관 여부를 확인하였다. 그리고 직접 확산 네트워크를 바탕으로 실제 발병한 12개 농가 각각의 발병 경로를 추적했고 직접 확산 네트워크로 추적이 어려운 농가들은 간접 확산 네트워크를 이용해서 발병 경로를 해석하였다. 그리고 약품 네트워크의 중심성 분석 3개, 사료 네트워크의 중심성 분석 3개, 공기 영향 네트워크의 중심성 분석 1개 등 총 7

개의 중심성 분석 가운데 감염을 가장 정확히 모의하는 요인을 분석하기 위해서 통계적 방법으로 이분형 로지스틱 분석을 수행하였다. 이분형 로지스틱 분석은 각각의 중심성 분석 결과와 감염 여부의 상관도를 나타내는 것으로 SPSS 20.0을 이용해서 수행하였다.

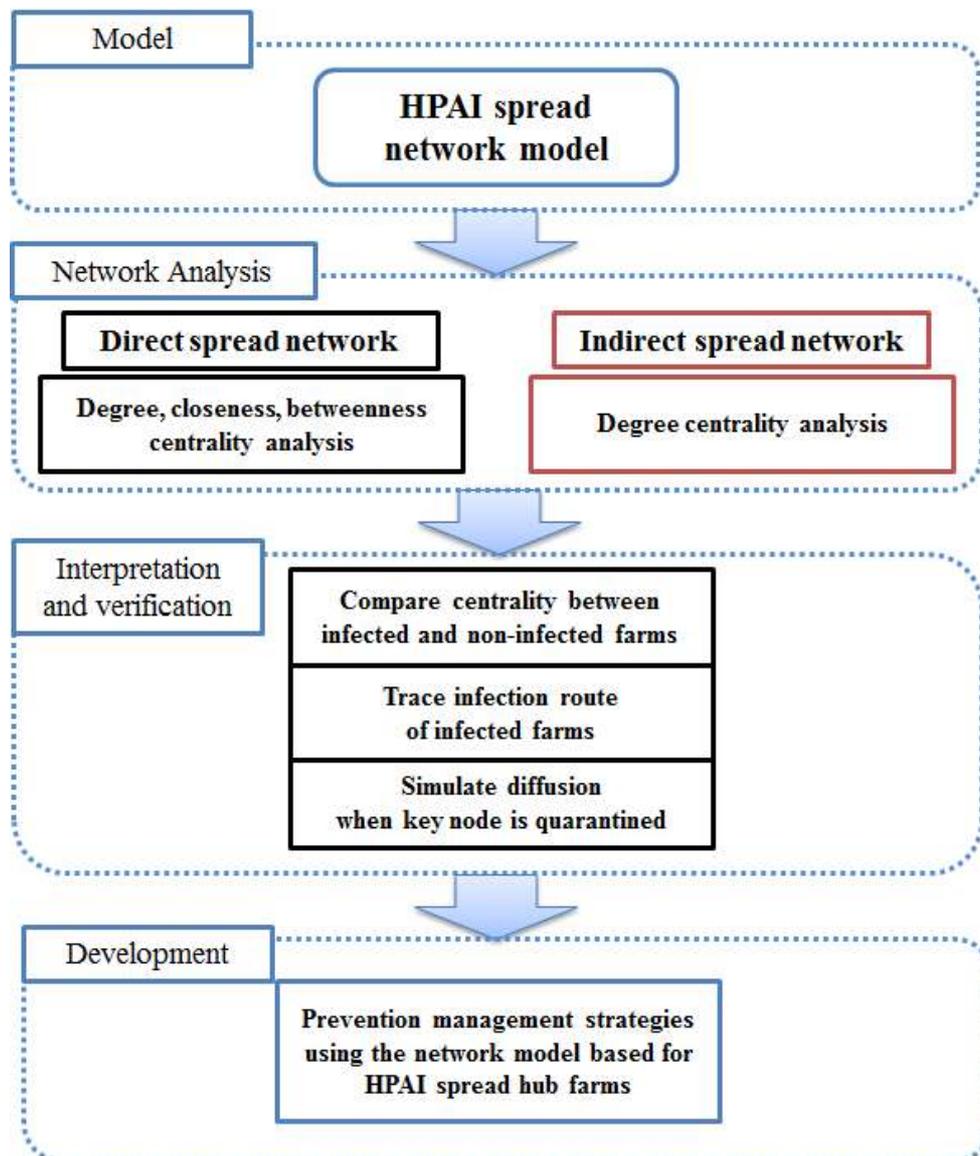


Fig. 3.5 Flow chart of the study procedure; Taking the 2008 outbreak case into consideration, effective measures were proposed to prevent HPAI from spreading at the earliest stages.

그 중 가장 큰 상관도를 가지는 분석 자료를 이용해서 주요 절점을 선정하지 않고 이를 확산으로부터 통제할 경우 농가로 확산을 얼마나 차단할 수 있는지를 모의하였다. 발병 이전에 통제하는 농가의 수가 많으면 당연히 비발병 농가의 수도 많아진다. 그러나 차단하는 농가의 수가 많아질수록 많은 비용이 수반되기 때문에 가장 효율적인 통제 농가수를 선정할 필요가 있다. 따라서 식 (5)와 같이 비발병 농가수에 차단농가수를 나눈 발병차단지수를 이용해서 효율적인 방역 체계를 위한 최적의 차단 농가수를 제안하였다.

$$Cost = \frac{\text{number of } QF}{\text{number of } AF}$$

$$Benefit = \frac{\text{number of } SF}{\text{number of } AF - \text{number of } QF} \quad (5)$$

$$PI = \frac{\text{number of } SF \times \text{number of } AF}{(\text{number of } AF - \text{number of } QF) \times \text{number of } QF}$$

*AF*: the total number of farms

*SF*: the total number of suspected farms

*QF*: the total number of quarantine farms

*PI* : Outbreak prevention index

### 3) 취약성 분석

연결 중심성, 근접 중심성, 매개 중심성과 같은 중심성 지수는 네트워크 내에서 질병 전파의 주요 역할을 하는 절점들을 찾아낼 수 있으므로 초기 확산에 대처하는데 주요한 역할을 한다 (Mansley et al., 2003). 중심성 지수가 높으면 주요 허브 지점에 절점들이 위치할 확률이 높기 때문에 감염에 취약하거나 다른 농가들에게 미치는 영향

력이 높다. 따라서 중심성 지수가 높은 절점을 우선적으로 관리할 때 네트워크 내의 다른 절점으로 확산을 차단하거나 그 속도를 늦출 수 있다. 2008년 HPAI 최초 발병지역인 전라북도 김제시 용지면의 39개 농가를 대상으로 가금 관련 업체 자료를 바탕으로 동물 약품, 가축 사료에 대한 직접 확산 네트워크, 공기중 영향도를 이용해서 간접 확산 네트워크를 구성하였고 중심성 분석을 통해서 HPAI 질병 확산 중요지점을 선정하였다. Table 1은 각각의 중심성 분석에서 평균값과 비교한 값이 높은 상위 5개의 절점 번호와 괄호 안은 그 값을 나타낸다.

사료 네트워크에서 3개의 중심성 지수에 대한 전체 농가 평균은 각각 0.240, 0.421, 0.391 이고 표준편차는 0.096, 0.067, 0.034 이다. 약품 네트워크와 같이 근접 중심성 지수의 평균이 가장 높았고 표준편차는 연결 중심성 지수가 가장 높았다. 중심성 지수가 높은 농가는 역시 다양한 가축사료업체를 이용하는 농가이거나 다양한 농가를 방문하는 업체를 이용하는 농가였다. 15번, 6번, 23번 농가는 근접 중심성에서, 15번 농가는 근접 중심성 지수에서, 6번 농가는 매개 중심성에서 가장 큰 값을 기록하였다. 그리고 근접 중심성에서 2번째로 높은 수치를 보인 13번 농가는 가장 높은 15번 농가와도 큰 차이가 나지 않았으며 타 농가들과 비교해 다른 측정 방법에서도 모두 높은 값을 보였다. 이들은 모두 2개 이상의 사료업체를 이용하거나 가장 많은 농가를 방문하는 사료 업체를 이용하고 있다.

동물약품과 가축사료 업체의 차량 및 관계자 등을 통한 HPAI 바이러스의 간접 확산을 방지하기 위해서는 밀집된 지역일 경우 가금 관련업체의 이용을 통제하는 방법으로 질병의 확산 위험을 예방할 수 있을 것으로 판단된다. 관련 제품의 출입 차량 및 관계자의 농가 방문에 대한 철저한 관리는 발병 이후 확산을 차단할 수 있을 뿐만 아니라 큰 피해를 사전에 효과적으로 차단할 수 있다.

Table 3.1 Centrality in HPAI diffusion network

Rank	Medicine Network			Feed Network			Air Network
	Degree	Closeness	Betweenness	Degree	Closeness	Betweenness	Degree
1	35(0.342)	29(0.507)	31(0.192)	15(0.342)	15(0.521)	6(0.148)	6(0.063)
2	7(0.342)	31(0.507)	29(0.130)	6(0.342)	13(0.514)	15(0.117)	10(0.060)
3	8(0.316)	7(0.500)	5(0.128)	23(0.342)	3(0.514)	38(0.100)	17(0.052)
4	2(0.316)	35(0.494)	7(0.121)	17(0.316)	1(0.514)	23(0.089)	31(0.044)
5	29(0.316)	8(0.487)	19(0.118)	10(0.316)	6(0.507)	17(0.073)	21(0.039)
avg	0.216	0.417	0.039	0.240	0.421	0.039	0.026
std.	0.085	0.056	0.051	0.096	0.067	0.034	0.014

Where avg is the average value of centrality value; std is the standard deviation of centrality.

네트워크 중심성 이론을 적용한 HPAI 확산 모델을 검증하기 위해 실제 발병농가와 비발병농가의 중심성 지수를 비교하면 Table 2와 같다. 직접, 간접 네트워크의 중심성 분석에서 감염농가는 비감염농가와 비교할 때 모든 지수가 높게 나왔다. 특히 약품 네트워크의 연결 중심성 분석에서 감염농가의 중심성 지수는 비감염농가의 중심성 지수보다 2.1배가 크며 매개 중심성 분석에서는 모든 감염농가가 전체 농가의 평균보다 높은 지수를 기록했고 감염농가의 중심성 지수가 비감염농가의 중심성 지수보다 1.55배 이다. 중심성이 높은 농가들이 직·간접적으로 바이러스에 취약한 위치에 있음을 확인할 수 있고 직·간접 확산 네트워크에서 중심성을 통한 감염농가 추정이 유효하다는 것을 확인할 수 있다.

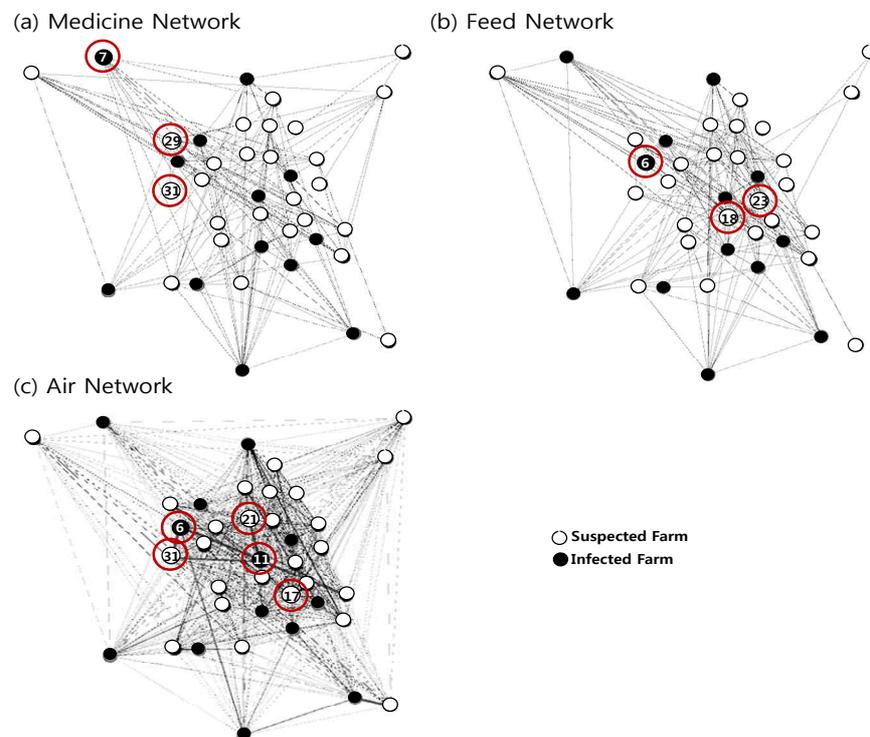


Fig. 3.6 Centrality analysis results from HPAI spread network: Red circles represent hub farms in a network that has a high value in centrality.

Table 3.2 Centrality infected/suspected farms

	Medicine			Feed			Air
	Network			Network			Network
	Degree	Closeness	Betweenness	Degree	Closeness	Betweenness	Degree
IP	0.252	0.445	0.051	0.254	0.431	0.043	0.027
NIP	0.200	0.405	0.033	0.234	0.417	0.038	0.025
IP/NIP	1.262	1.101	1.534	1.088	1.034	1.134	1.074

다양한 가금관련업체를 이용하는 농가는 그렇지 않은 농가보다 HPAI 감염 확률이 상대적으로 높아질 수 있고 많은 농가를 고객으로 가지는 가금관련업체가 방문하는 농가 역시 HPAI 감염에 취약할 수 있다. 그리고 기하학적으로 주변 농가와 밀접하고 도로와 인접한 지역에 위치한 농가 역시 공기를 통한 바이러스의 간접적 확산에 노출될 가능성이 높다.

공기 영향 네트워크에서는 농가간의 공기중 영향도에 따라서 중심성 분석을 수행한 것으로 대상지역에서 농가들이 밀집된 중심 지역에 위치한 농가들이 상대적으로 degree 중심성 지수가 높았다. 전체 농가의 평균이 0.256인데 상위 5개 농가들은 평균을 훨씬 상회하는 값을 가지고 있으며 분산 역시 0.142로 다른 네트워크의 지수들과 비교해서 높은 값을 보였다. 그리고 상위 5개 농가의 위치는 약품 네트워크, 사료 네트워크에서 높은 중심성 지수를 기록한 농가들과 달리 모두 기하학적으로 밀집한 중심 지역에 있으며 도로에 아주 인접한 위치에 있다. 밀집된 지역의 농가가 감염될 경우 약품 및 사료 업체 이용과 무관하게 공기중 영향에 따라 인접한 농가들에게 바이러스가 확산된다. 그리고 도로에 인접해 있으면 바이러스를 가진 차량에 의한 간접적 확산의 영향도 크다. 따라서 밀집된 지역이나 도로에 인접한 농가들은 HPAI 발병을 차단하기 위해서는 가금관련업체 이용에 관해서 주의를 기울여야할 뿐만 아니라 공기중 확산을 방지하기 위한 대책이 필요하다.

#### 4) 발병 경로 추적

2008년 김제시 농가에서는 총 12개의 농가가 약 2주간에 걸쳐서 1번 농가부터 12번 농가까지 순차적으로 발병되었으나 역학조사만으로는 발병 경로를 추적하기 어렵다. 먼저 약품 및 사료 업체의 이용에 따른 간접 확산 네트워크를 이용해서 발병 농가의 발병 경로를 분석해보고자 한다.

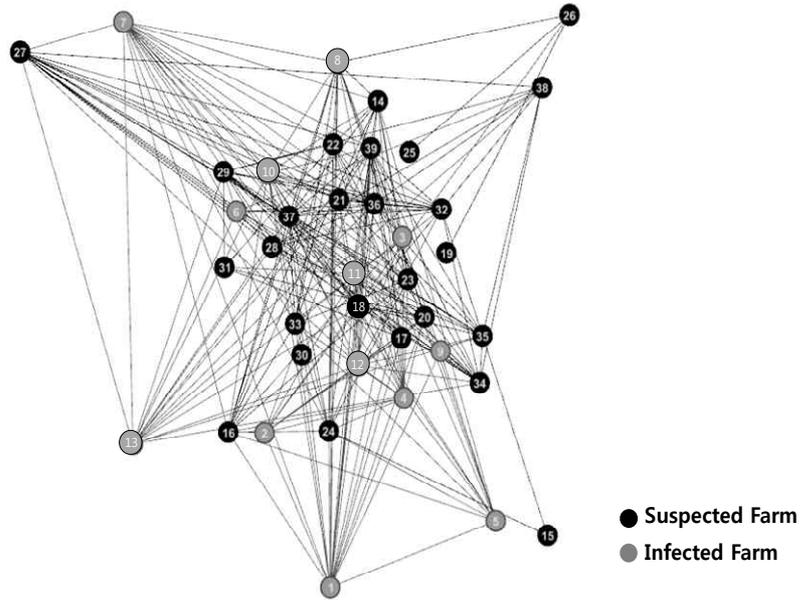


Fig. 3.7 The direct HPAI spread network with two factors, animal feed and medicine

Fig. 3.7의 간접 HPAI 확산 네트워크는 가축사료, 동물약품의 2가지 요인으로 구성된다. 발병 당시 가축사료, 동물약품을 제외한 다른 자금 관련 업체의 이동이 없었던 것으로 조사되어 두 가지의 직접적 확산 요인만을 이용해서 당시 발병 상황을 추적하고자 한다. 최초 발병 농가와 다른 농가와의 연결선을 통해 연결단계를 구분할 수 있다.

현재 방역 대책과 같이 공기를 통한 직간접 확산의 가능성을 무시하고 매개체를 이용한 직접적인 접촉에 의한 HPAI 질병 확산만으로 추적한다면 농가들은 최초 발병 농가인 1번 농가와 1단계로 접촉한 농가, 타농가를 거쳐 2단계로 접촉한 농가로 구분할 수 있다. 인접 행렬(adjacent matrix)에서 최초 발병 절점을 선택하고 그와 연결된 절점을 차례대로 검색한다. 그리고 연결된 절점을 그룹화한다. 그리고 그 그룹의 절점과 연결된 절점을 각각 하위 그룹에 포함시킨다. 그러면 최초 발병 절점과 단계별로 연결된 그룹이 구성된다. 하위 그룹을 형성하는 과정을 상위

그룹에 포함된 절점의 수만큼 반복하면 그만큼의 경로 수가 생겨나고 Fig. 3.8과 같은 감염 경로를 구축할 수 있다. 추후에 감염 확률과 같은 자료가 추가되더라도 자동화되어있는 Fig. 3.9의 과정에 따라서 손쉽게 감염 경로를 구축할 수 있다. 최초 발병 농가인 1번 농가의 발병을 가정하고 HPAI 질병 확산을 모의했을 때 네트워크 연결선을 이용해서 도식화하면 Fig. 3.8과 같다. 약품 네트워크와 사료 네트워크를 합한 직접적 확산 네트워크는 총 39개의 절점과 372개의 연결선을 가진 가중 네트워크로 절점당 평균 9.54개의 연결선을 가진다.

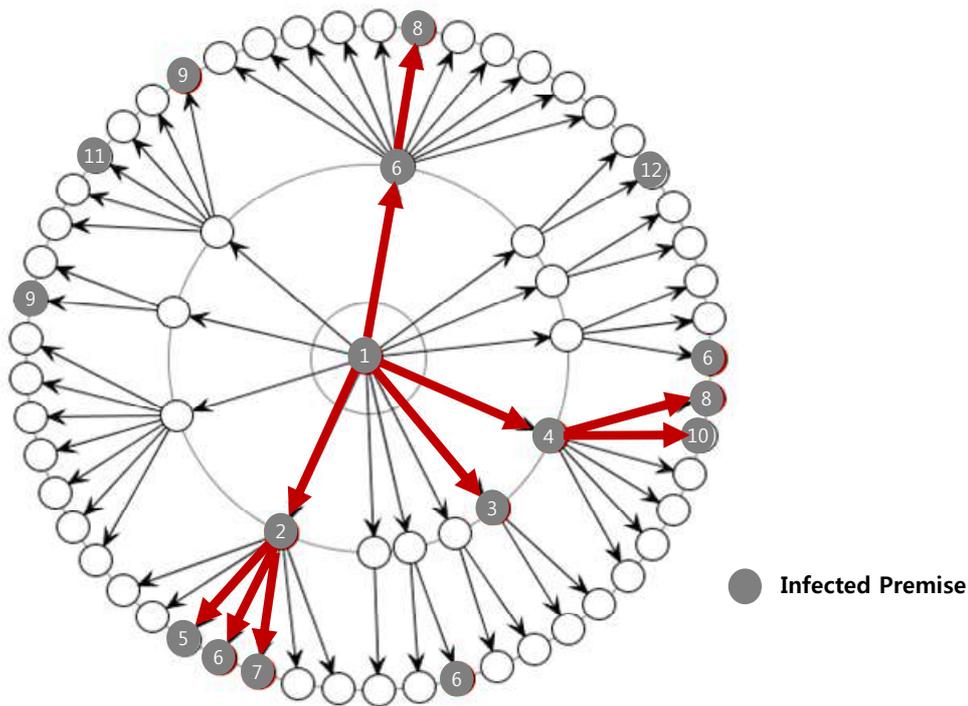


Fig. 3.8 Trace the infection route diagram: This study could simulate the spread pattern using the direct HPAI spread network model.

Fig. 3.8을 보면 발병 농가 중 2, 3, 4, 6번 농가는 1번 농가와 1단계로 연결되어 있으며 5, 6, 7, 8, 9, 10, 11, 12번 농가는 2단계로 연결되어 있다. 이 중 6번 농가는 1번 농가와 직접적으로도 연결되어 있으며 2번 농가를 거쳐

서도 연결되어 있음을 확인할 수 있다. 바이러스의 확산 과정을 발병 농가 경로에 따라 추적해보면 9, 11, 12번 농가를 제외한 다른 농가들은 모두 매개체를 통한 직접 확산의 발병 경로를 확인할 수 있다. 그리고 시기별로 2, 3, 4번 농가 발병 이후 3일정도의 시간이 지난 후 다른 농가들이 발병한 것을 통해서도 Fig 4의 단계별 확산 과정을 설명할 수 있다. 다만 9, 11, 12번 농가는 1번 농가와 연결되어 있으나 1단계로 연결되어 있지 않으며 발병 농가를 거쳐서는 연결되어 있지 않기 때문에 직접 확산 요인만을 이용한 네트워크로는 발병 경로를 설명하기 힘들다. 3개의 농가 가운데 9, 11번 농가는 축산농가가 밀집한 지역에 위치하고 있다. 간접 확산의 영향을 고려한 공기 영향 네트워크에서 9번 농가와 11번 농가의 연결 중심성 지수는 각각 0.350, 0.313으로 전체 농가의 평균인 0.269와 비교해서 상당히 높은 값을 보인다. 그리고 앞서 발병한 2, 4 농가로부터 대상 지역의 주풍방향인 서풍의 영향에 따라 큰 영향을 받은 것을 확인할 수 있으며 발병 경로도 이를 통해서 설명할 수 있다. 단, 12번 농가의 발병 경로는 직접 확산 네트워크로 설명할 수 없으며 공기 영향 네트워크에서 연결 중심성 지수도 0.209로 전체 농가의 평균에 미치지 못한다. 다만 12번 농가는 대상 지역에서 인근 지역의 가금 관련 대기업으로 연결되는 차량의 통행이 많은 도로와 인접한 위치에 있는데 바이러스를 가진 차량을 통한 공기중 확산의 영향을 받을 가능성이 높다. 그렇지 않다면 현재 자료로 구성된 HPAI 확산 요인 외에 다른 요인에 의한 전파일 가능성이 높다. 역학조사를 통한 자료의 질적 양적 발전이 이루어진다면 보다 정확한 경로 추적이 가능할 것으로 판단된다.

총 39개 농장 가운데 12개의 발생 농가 중 9개 발생 농장의 경로 추적을 매개체를 이용한 직접 확산 경로로 확인했으며 나머지 3개 농가 중 2개 농가는 공기중 간접 확산의 영향도가 큰 농가로 확인하였다. 확산 요인에 대한 추가적인 자료 수집이 가능하면 확산 네트워크를 추가할 수 있을 것이며 보다 정확도가 높은 HPAI 질병 확산 통합 네트워크 구성이 가능할 것이다. 그리고 중심성 지수를 통한 보다 정확한 확산 중요 지점 선정을 할 수 있을 것으로 판단된다.

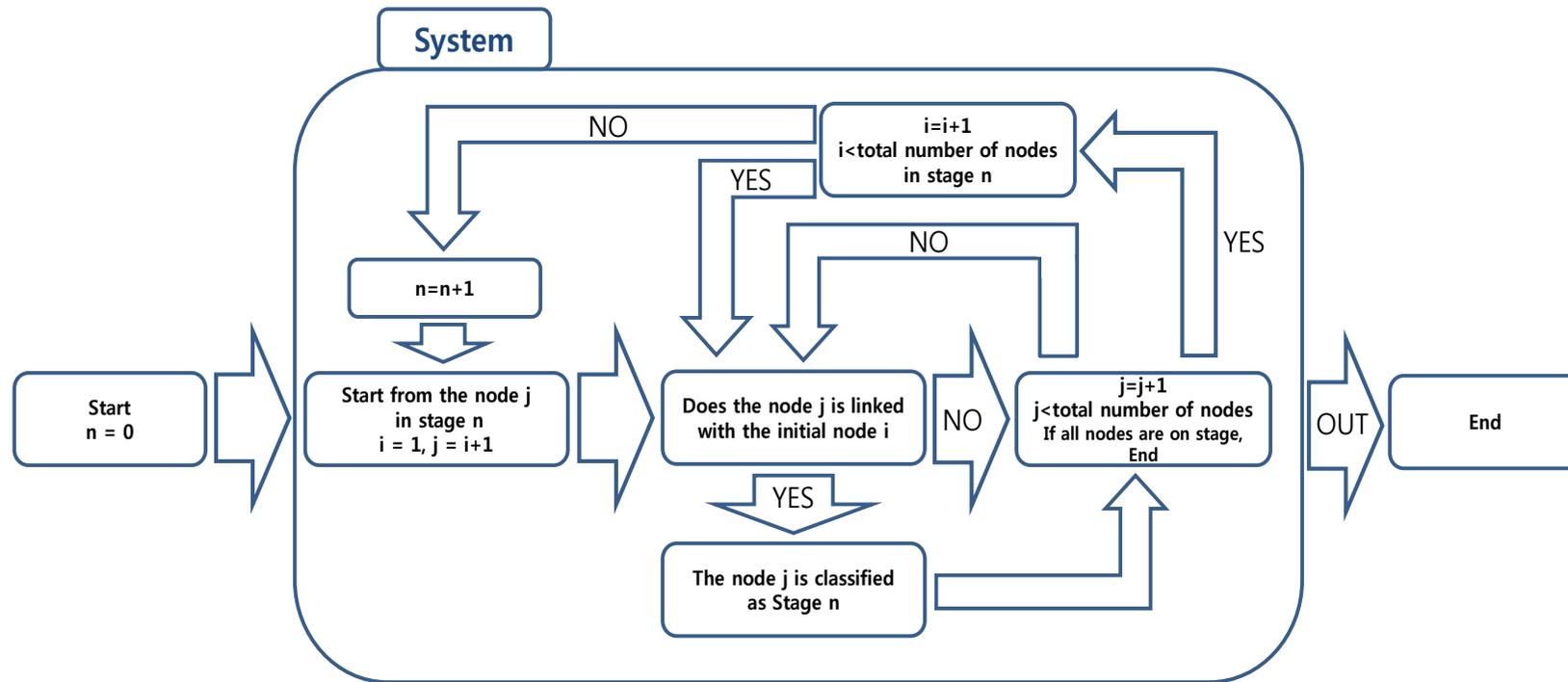


Fig. 3.9 System flow of constructing the infection route diagram based on adjacent matrix data

## 5) 절점 차단 모의

확산 차단을 위해 네트워크 해석으로 중심성 지수가 높은 주요 절점을 선정하여 질병 발생 전에 확산 요인을 통제 대책을 집중하여 네트워크 내 고립도를 높여 전체 구조의 확산을 막을 수 있다. 직접, 간접 확산 네트워크에서 중심성 지수가 높은 농가에 한해 평소 철저한 방역 관리를 시행하거나 다른 농가와 의 이용도가 높은 업체 대신에 다른 업체의 이용을 권장하는 방식으로 감염 확률을 낮출 수 있다. 중심성이 높은 농가는 감염에 취약하므로 타농가로 바이러스를 확산시키는 영향도 역시 높기 때문에 주요 농가를 관리 대상으로 지정하여 예방 대책을 집중한다면 임의의 농가에서 발병해서 확산되더라도 지역내 감염 농가의 수를 줄일 수 있다.

약품 네트워크의 중심성 지수 3개, 사료 네트워크의 중심성 지수 3개, 공기 영향 네트워크의 중심성 지수 1개 등 총 7개의 요인별 중요도 분석을 위해 이분형 로지스틱 모형을 이용하였다. 감염 여부를 종속변수로 두고 나머지 7개의 요인을 공변량으로 통계적 분석을 수행했을 때 결과는 Table 3과 같다. 7가지 요인의 유의확률은 0.05보다 낮은 적절한 수준으로 평가되었고 약품 네트워크의 사이 중심성 지수의 B가 14.019로 감염에 가장 큰 영향을 미친 요인으로 분석되며 공기 영향 네트워크의 연결 중심성 지수의 B는 1.256으로 가장 영향을 적게 미친 요인으로 나타난다. 따라서 약품 네트워크의 사이 중심성 지수가 높은 상위 3개 농가, 5번, 29번, 31번 농가를 차단했을 경우를 차단하지 않은 경우와 비교하면 Fig. 10과 같다. 차단하지 않은 경우는 확산이 진행되면서 전체 농가의 77%, 총 30개의 농가가 바이러스의 영향을 받을 수 있는 상황이 되나 중심성이 높은 3개의 절점을 차단하면 차단한 농가를 포함해서 총 20개 농가가 바이러스의 영향을 받지 않고 단지 19개인 49%의 농가만이 감염의 위험에 노출된다. 확산 단계가 지날수록 그 차이는 더욱 커질 것

으로 예상된다. 그리고 최초 발생 농가 이외의 지역으로 확산될 경우 네트워크의 규모는 더욱 커질 것이며 확산 차단을 위한 주요 농가 통제 효과는 훨씬 클 것으로 기대된다.

중심성이 높은 농가를 통제하는 수를 늘리면 발병 농가의 수도 역시 줄일 수 있다. 그러나 발병 이전에 중심성이 높은 농가의 확산을 완전 통제하기 위해서는 상당한 경제적 비용이 든다. 따라서 가장 효율이 높은 방역 대책을 위해서 중심성이 높은 농가의 적정 차단수를 결정할 필요가 있다. 약품 네트워크에서 사이 중심성 지수가 높은 농가의 통제 정도를 조절하여 확산 양상을 모의하고자 한다.

최초 발병 농가인 1번 농가에서 발병했을 때 통제된 농가 수와 비발병 농가 수의 비를 이용해서 발병차단지수를 나타낼 수 있다. 발병차단지수는 통제 농가당 비발병 농가수를 계산하는 것으로 값이 높을 수록 더 효율적인 통제가 가능하다는 것을 보여준다. Table 4는 전체 농가 중 사이 중심성 지수가 높은 농가, 1%, 5%, 10%를 차단했을 경우에 각각 얼마나 효율적으로 확산을 차단할 수 있는지를 보여준다. 총 39개 농가 중 지수가 가장 높은 상위 1%, 1개 농가를 통제했을 경우 비발병 농가 수는 7개 농가이다. 발병 차단 지수는 비발병 농가 수에 차단 농가 수 1을 나눈 7이 된다. 5%인 2개 농가를 차단할 경우 발병 차단 지수는 9로 1개 농가를 차단했을 때의 7보다 높게 나왔다. 그리고 10%를 통제했을 경우는 5.3으로 5%를 통제했을 때보다 오히려 낮은 값이 나왔다. 통제하는 농가의 수가 많아질수록 확산 차단의 효과는 점차 커지지만 차단 효율은 5%를 통제했을 때가 1%, 10%를 통제했을 때보다 가장 높은 것을 확인할 수 있다.

Table 3.3 Factor analysis using a binary logistic model

		B	S.E.	Sig.
Medicine Network	degree	2.211	3.025	0.001
	closeness	4.652	0.556	0.005
	betweenness	14.019	2.055	0.008
Feed Network	degree	5.116	1.150	0.007
	closeness	3.287	0.341	0.003
	betweenness	7.969	1.249	0.001
Air Network	degree	1.256	1.728	0.005

B is the estimation of logistic regression coefficients; S.E is the standard error; Sig. is the significance probability.

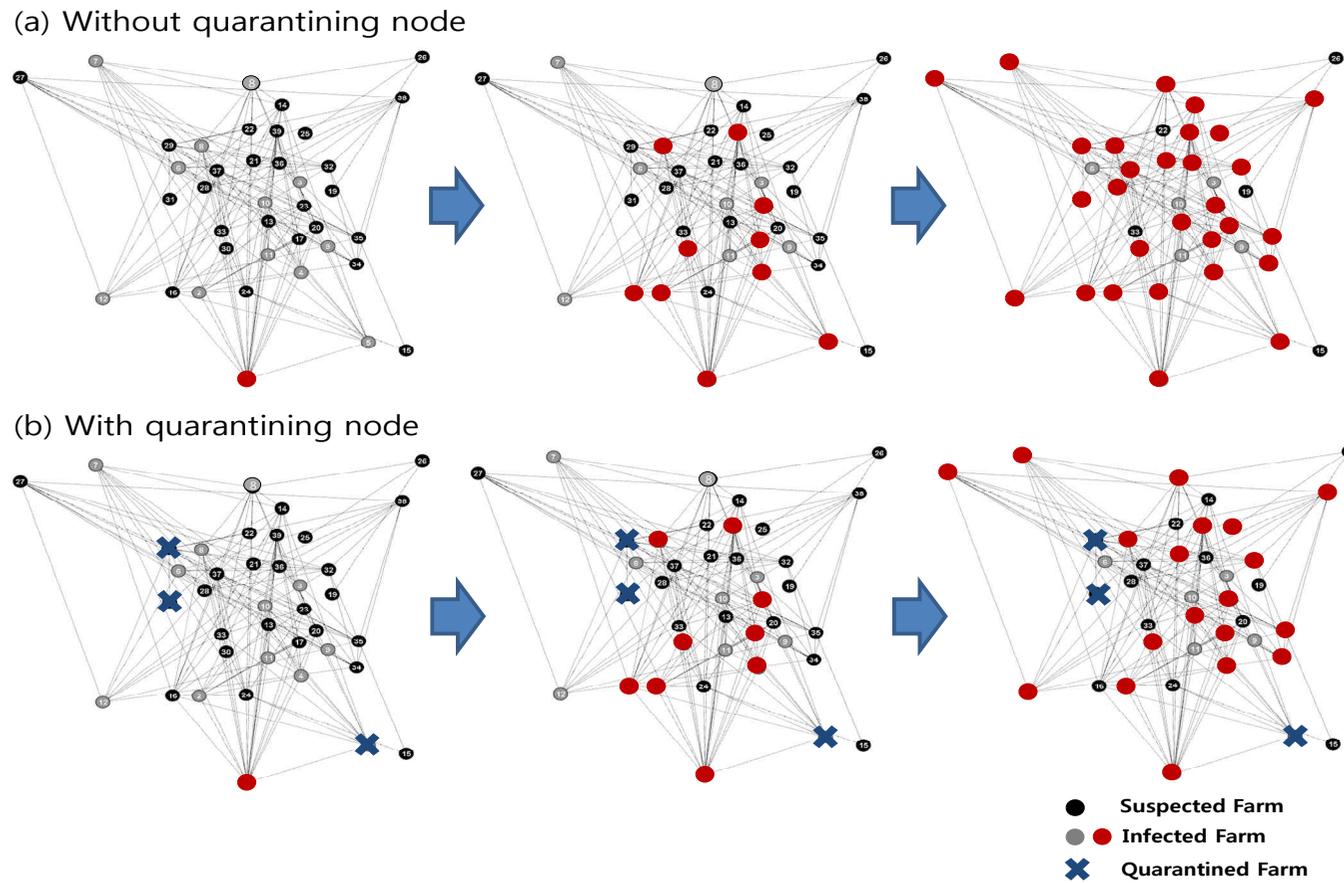


Fig. 3.10 Comparison of the tendency of HPAI spread with/without quarantining farm  
 Check point X is a quarantined farm. Grey circle are the suspected farms, black circles,  
 and red circles are the infected farm.

Table 3.4 Outbreak prevention index with quarantined farm

	Top 1% quarantined	Top 5% quarantined	Top 10% quarantined
# <i>SF</i>	7	18	21
# <i>QF</i>	1	2	4
<i>PI</i>	7	9	5.3

*PI* : Outbreak prevention index

*SF*: the total number of suspected farms

*QF*: the total number of quarantine farms

# of suspected farms is the number of suspected farms; # of quarantined farms is the number of quarantined farms.

본 연구에서 수행한 총 7가지 분석 가운데 약품 업체 이용 자료를 기초로 구성된 약품 네트워크의 사이 중심성 지수가 감염에 가장 큰 영향을 미치는 것으로 나오며 전체 농가 중 중심성 지수가 높은 농가, 5%를 차단했을 때 발병 차단 효율이 가장 높은 것을 확인할 수 있다. 확산 단계가 지날수록 감염 농가 수는 기하급수적으로 증가한다. 실제 발병시 소지역 단위를 벗어나 전국적으로 확산될 경우 초기 대응의 차이에 따라서 발병 농가의 규모는 더욱 커질 것이다. 따라서 예방을 위해 중심성이 높은 일부 농가를 우선적으로 차단하는 방법은 효율적인 방역대책이 될 수 있다.

### 3.5 소 결

허브 네트워크에 존재하는 허브 절점은 취약성 분석에 주요한 역할을 한다. 허브 절점에 대한 집중적인 공격은 연결성의 대부분을 차지하고 있어 전체 네트워크 구조를 붕괴시킬 수 있다. 따라서 허브 네트워크에

서 취약성을 분석은 곧 절점들의 중요도를 정량화하는 것이다. 본 연구에서는 취약성 평가를 위해서 중심성 분석을 수행했으며 HPAI 확산 사례에 적용하고자 하였다.

HPAI는 확산 속도도 빠르고 잠복기가 존재하기 때문에 발병을 확인하고 농가 및 일정 반경 내의 농가를 폐쇄하는 등의 방역 대책을 수행하면 이미 질병이 확산된 이후이므로 발병 경로를 추적하고 차단하는데 문제가 생길 가능성이 크다. 그리고 연구의 대상지역과 같이 가금 농가가 밀집된 지역은 동종 가금 업체를 이용하고 공기 중 확산이 용이해서 전파 요인이 복잡하게 얽혀있다. 따라서 임의의 한 농가에서 HPAI가 발병하면 관계되는 다른 농가들이 빠른 시간 내에 연쇄적으로 감염될 가능성이 크다. 가금농가가 집단화된 밀집 지역의 경우 다른 지역과도 다양하게 연계되어있으므로 초기 확산 대응을 실패하면 전국적으로 확산되어 피해 규모가 더욱 커진다. 이 경우에 밀집 지역 농가의 가금업체이용정보와 같은 기존 출입 상황 자료와 공기중 확산 자료를 이용한 네트워크 중심성 분석을 통해 질병에 취약한 농가를 산정할 수 있고 예방 대책을 집중할 수 있어 효율적으로 피해 규모를 줄일 수 있다.

본 연구는 주기적으로 농가를 방문하는 가금 관련 업체를 매개체로 이용하는 바이러스의 간접 확산과 공기를 통해 이동하는 바이러스의 직접 확산을 포괄적으로 해석하기 위해서 가금 농가간의 HPAI 확산 네트워크를 이용하였다. 가금 이용 자료를 바탕으로 구축한 네트워크는 연결선 수 분포가 멱급수의 법칙을 따르는 허브 네트워크이기 때문에 허브 농가의 영향력이 상당히 크다. 동일 업체의 이용을 기반으로 구성된 직접 확산 네트워크와 공기 이동에 따른 간접 확산 네트워크에서 중심성 분석을 수행한 결과 실제 발병농가들은 비발병농가와 비교했을 때 중심성 지수가 더 높은 것으로 나왔으며 발병 농가의 바이러스 확산 경로도 직, 간접적 요인을 통합적으로 이용해서 해석할 수 있었다. 중심성 지수가 큰 농가는 감염에 취약할 뿐만 아니라 전체 네트워크로 확산시킬 수 있는

과급력 또한 크다. 따라서 HPAI 확산 네트워크에서 중심성이 높은 농가에 방역 대책을 집중하여 관리할 필요가 있다. 통계적 분석을 이용해서 가장 영향도가 높은 중심성 분석 자료를 선정했으며 효율적인 방역대책 수립을 위한 적정 통제 농가 수를 제안하였다. 본 연구 결과를 가금 산업 농가가 밀집된 지역에 확대 적용했을 때 효율적인 예찰 시스템의 기초 자료로 구성할 수 있을 것이며 발생 이후에도 확산 경로를 추적하는데 주요한 역할을 담당할 수 있을 것으로 판단된다. 그리고 현재 수집된 자료 이외에 추가적인 확산 요인의 조사가 수행되면 보다 신뢰성 높은 HPAI 확산 네트워크를 구성할 수 있을 것으로 사료된다.

## 참고문헌

1. Albert, R., H. Jeong, and A.-L. Baraba'si, 1999. Diameter of the world-wide web. *Nature* 401: 130-131.
2. Albert, R., H. Jeong and A.-L. Baraba'si, 2000. Attack and Error Tloerance of complex network. *nature* 406: 378-382.
3. Aldana, M. and P. Cluzel, 2003. A natural class of robust networks, *Proceedings of the National Academy of Sciences* 100(15): 8710-8714.
4. Animal and Plant Quarantine Agency, 2008. Republic of Korea, 2008 HPAI Epidemiological investigation report. <http://naver.nanet.go.kr/MarcView.do?cn=MONO1200832687&query=&zone>.
5. Bearman, P. S., J. Moody and K., Stovel, 2004. Chains of affection: The structure of adolescent romantic and sexual networks1. *American Journal of Sociology* 110(1): 44-91.
6. Beauchamp, M. A., 1965. An improved index of centrality. *Behavioral Science* 10(2): 161-163.
7. Boccaletti, S., V. Latora, Y. Moreno, M. Chavez and D. U. Hwang, 2006, Complex networks: Structure and dynamics. *Physics Reports* 424(4): 175-308.
8. Callaway, D. S., M. E. J. Newman, S. H. Strogatz and D. J. Watts, 2000. Network robustness and fragility: Percolation on random graphs. *Physical review letters* 85(25): 5468-5471.
9. Cohen, R., K. Erez, D. Ben-Avraham and S. Havlin, 2000. Resilience of the Internet to random breakdowns. *Physical review letters* 85(21): 4626 - 4628.

10. Cohen, R., K. Erez, D. Ben-Avraham and S. Havlin, 2001. Breakdown of the Internet under intentional attack. *Physical Review Letters* 86(16): 3682–3685.
11. Dorogovtsev, S. N., A. V. Goltsev and J. F. Mendes, 2008. Critical phenomena in complex networks. *Reviews of Modern Physics*. 80(4): 1275–1335.
12. Dunne, J. A., R. J. Williams and N. D. Martinez, 2002. Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters* 5: 558–567.
13. Fell, D. A. and Wagner, A., 2000. The small world of metabolism. *Nature Biotechnology* 18(11): 1121–1122.
14. Freeman L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239.
15. Gupta, S., R. M. Anderson and R. M. May, 1989. Networks of sexual contacts: Implications for the pattern of spread of HIV. *AIDS* 3(12): 807–817.
16. Huberman, B. A., 2001. *The Laws of the Web: Patterns in the ecology of information*. MIT Press.
17. Jeong, H., S. P. Mason, A.-L. Barabási and Z. N. Oltvai, 2001. Lethality and centrality in protein networks. *Nature* 411: 41–42.
18. Klovdahl, A. S., J. J., Potterat, D. E., Woodhouse, J. B., Muth, S. Q., Muth and W. W., Darrow, 1994. Social networks and infectious disease: The Colorado Springs study. *Social Science & Medicine* 38(1): 79–88.
19. Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley and Y. V. Aberg, 2001. The web of human sexual contacts. *Nature* 411: 907–908.
20. Loomis, C. P. and C. Proctor, 1950. The relationship between

- choice status and economic status in social systems. *Sociometry* 13(4): 307–313.
21. Mansley, L. M., P. J. Dunlop, S. M. Whiteside and R. G. H. Smith, 2003. Early dissemination of foot-and-mouth disease virus through sheep marketing in February 2001. *The Journal of veterinary record* 153(2): 43–50.
  22. Martinez M., A. M. Perez, A. Delatorre, I. Iglesias, J. Sanchez-Vizcaino and M. J. Munoz, 2011. Evaluating surveillance in wild birds by the application of risk assessment of avian influenza introduction into Spain. *Epidemiology and Infection* 139(1): 91–98.
  23. May, R. M. and R. M. Anderson, 1989. The transmission dynamics of human immunodeficiency virus(HIV). *Applied Mathematical Ecology* 263–311.
  24. Moreno, Y., M. Nekovee, and A. F. Pacheco, 2004. Dynamics of rumor spreading in complex networks. *Physical Review* 69(6): 066130.
  25. Nam, V., World Health Organization, H5N1 avian influenza: Timeline of major events. [http://www.who.int/influenza/human\\_animal\\_interface/avian\\_influenza/H5N1\\_avian\\_influenza\\_update.pdf](http://www.who.int/influenza/human_animal_interface/avian_influenza/H5N1_avian_influenza_update.pdf). 15 June. 2012.
  26. Newman, M. E. J., 2002. Spread of epidemic disease on networks. *Physical Review* 66(1): 016128.
  27. Sole, R. V. and M. Montoya, 2001. Complexity and fragility in ecological networks, *Proceedings of the Royal Society of London. Biological Sciences* 268(1480): 2039–2045.
  28. Tu, Y., 2000. How Robust is the Internet?. *Nature* 406: 353–354.
  29. White, J. G., E., Southgate, J. N., Thompson and S., Brenner,

1986. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 314(1165): 1-340.
30. Woo, B. J., H. W. Lee, Y. J. Hwang, J. M. Lee and J. N. Kim, 2008. Economic Impact of Avian Influenza outbreak. Research Report of Korea Rural Economic Institute p.102.
31. World Health Organization, Area with confirmed human cases of H5N1 avian influenza since 2003. <http://reliefweb.int/sites/reliefweb.int/files/resources/A724B3AD85DDDF2C852577230060879A-map.pdf>. 9 May. 2010.

## 제 4 장 확률을 이용한 네트워크 군집화 분석

### 4.1 서론

네트워크에서 공통된 특성을 가진 절점들이 군집되는 것은 자연스러운 일이다. 절점들은 다른 절점들과 관계, 전체 구조에서 위치 등에 따라서 특정 집단에 속하려는 성질을 가진다. 네트워크에서 어떤 절점이 어떤 소집단에 포함되어있는지에 대한 문제의 해석 결과는 절점들의 관계에 대한 이해를 도울 수 있을 뿐만 아니라 같은 집단의 다른 절점을 통해서 절점의 기능적인 특성을 예측할 수 있기 때문에 유용하게 활용될 수 있다. 절점은 자신과 밀접하게 연결된 절점들과 소집단을 형성하고 그 소집단의 정체성을 대변한다. 유사하게 소집단들이 모이면 전체 네트워크를 구성하기 때문에 전체 네트워크 해석은 소집단들의 정체성에 대한 분석에서 시작한다고 볼 수 있다. 군집화 방법을 적용하여 불균등한 허브 네트워크에서 소집단들을 구분하는 것은 네트워크의 특성을 해석하는데 주요한 방법이다.

네트워크 구조에서 소집단을 찾아내는 방법은 그래프 이론에서 큰 그래프를 작은 단위의 그래프로 분할하는 문제의 해법과 동일하다. 이 방법은 구조적 문제에서 하중 분할이나 네트워크 설계의 회로 분할 등에 다양하게 적용된다 (Pothén et al., 1990; Roy et al., 2003; Zhou, 2003). 또한 전체 구조의 특성이나 절점들을 연결하는 연결선의 특성에 따라 다양한 알고리즘을 이용하여 소집단을 구하는 방법론에 대한 연구가 활발히 진행되고 있다 (Girvan and Newman, 2002; Gudkov J. et al., 2002; Guimera et al., 2003; Leskovec et al., 2007, 2010; Marcos et al., 2008; Reichardt and Bornholdt 2004, 2006; Wu and Huberman, 2004). 그래프 이론을 기초로 네트워크 구조에서 소집단을 찾아내는 방법은 매우 유

용하나 절점과 연결선의 특성이 결정론적인 수치를 가지는 네트워크 문제에서 적용할 때만 유효하였다. 절점과 연결선을 대표하는 수치들이 확률론적 범위를 가질 때 복잡계 네트워크를 기반으로 한 네트워크 구조 및 주요한 절점들의 위상에 대한 해석을 시도한 연구(Fefferman, 1970; Guimera et al., 2003; Leskovec et al., 2010; Peter et al., 2010)나 절점을 확률변수로 하고 연결된 두 절점의 연계성을 나타내고 이를 이용하여 군집화를 시도하는 연구(Dechter and Pearl, 1989; Friedman et al., 1997; Friedman, 2004)가 있으나 연결선과 절점에 관한 확률 정보를 포괄적으로 동시에 분석하는데 한계가 있다.

확률론적 네트워크 해석의 보다 정확한 모의를 위해서는 연결선의 특성뿐만 아니라 절점의 특성도 고려할 필요가 있다. 네트워크 이론에서는 절점은 연결선과 관련된 정보만을 가지고 있으며 연결선은 시작 절점과 도착 절점 및 전달하는 값에 대한 정보를 가지고 있다 (Newman and Girvan, 2003). 주로 네트워크 구조는 한 가지 물질로 지배되는 단일 네트워크를 표현하고 절점을 통해서 운반되는 그 유일한 물질의 크기에 대한 정보는 연결선이 가지고 있으며 절점은 방향성만 결정한다. 그러나 절점은 운반되는 그 물질에 대한 저항적 특성을 가질 수 있다. 그리고 절점이 가지는 저항과 연결선이 가지는 값이 각각 확률적 분포를 가지고 있다면 확률적 연산에 이용되는 컨볼루션을 이용하여 이 둘을 연산하고 새로운 하나의 확률적 분포로 재규격화 할 수 있다.

본 연구에서는 절점들의 관계 정보들을 기반으로 네트워크 군집화에 대해서 해석하되 그 관계가 확률론적으로 구성되어 있을 때 수학 연산자인 컨볼루션을 활용하는 방법론을 개발하였고 이를 가상의 네트워크 구조에 적용하고 기존의 방법과 비교 검증하였다. 이 방법론은 절점과 연결선이 가지는 확률 분포 등의 정보를 포괄적으로 해석하여 네트워크 내부 변화의 고찰을 통해서 절점들을 군집화 할 수 있을 것이며 응력 흐름의 해석과 같은 구조 문제에 적용할 수 있을 것으로 기대된다.

## 4.2 확률 재규격화를 이용한 군집화

### 4.2.1 확률 재규격화

네트워크의 연결선이 가지는 값은 출발 절점에서 주어지는 값으로만 결정되지 않는다 (Newman and Girvan, 2003). 연결선의 값이 절점의 특징적인 저항력에 의해서 영향을 받기 때문에 절점으로 들어오는 외력과 절점이 가지고 있는 고유의 저항력, 즉 두 사상을 포괄적으로 해석할 수 있는 연산이 필요하다.

임의의 입력신호에 따라 출력이 결정되는 시스템에서 두 개의 독립된 신호에 관한 확률을 연산하기 위한 방법으로 컨볼루션(convolution)이라는 수학적 연산자를 이용한다 (O'Neil, 1963). 컨볼루션 기법을 통하여 두 개의 독립된 확률적 사상을 통합하여 새로운 사상으로 재해석하고 하나의 확률 사상으로 재규격화 할 수 있다. 그 기작은 입력값을 y축으로 대칭이동한 후에 단계별로 출력의 확률값을 곱해나가는 방식으로 연산을 수행해 나가는 방법이다 (Ringer, 1971; Burt and Garman, 1971). 시스템 입력신호의 확률분포를  $I(t)$ 라 하면 출력에 대한 확률적 연산에 의한 값은 특정시점  $t$ 에서의 입력신호를 기준으로 그보다 작은  $-\infty$ 까지 입력신호 값의 합과 출력신호,  $O(t)$ 의 곱으로 식 (1)과 같다.

$$y(x) = \left( \int_{-\infty}^t I(\tau) d\tau \right) \times O(t) \quad (1)$$

$I(t)$ : Input Signal

$O(t)$ : Output Signal

본 연구에서는 절점에서 주어지는 외력에 대한 저항력의 두 확률적 사상을 합성하여 그 연결선의 값을 구하는데 시스템에서 얻어지는 값을 이전의 입력에 대한 출력으로 해석하는 컨볼루션 기법을 적용하고자 한다.

Fig. 4.1과 같이 기존의 시스템의 입력과 출력에 대한 확률분포가  $y(x)$ 라는 새로운 확률 분포로 주어진다. 네트워크 모델에서 절점으로 들어오는 연결선이 확률분포를 가지는 외력(Load)을 가지고 있고 절점이 저항(Resistance)을 가지고 있다면 컨볼루션 연산을 통해서 절점을 나가는 연결선은 새로운 값( $W(x)$ )를 가지게 된다.

컨볼루션 연산은 두 확률 사상의 곱이라고 할 수 있다. 절점을 지나서 다음 절점으로 흘러나가는  $W(x)$ 를 구하기 위해서는 외력에서 저항을 뺀 값이 될 것이며 확률적 사상으로 대입하면 Fig. 4.1의 빗금친 영역으로 나타낼 수 있다. 그 영역을 구하기 위해서는 외력과 저항, 두 가지 확률 사상의 곱이 필요하다. 확률 사상의 곱으로 새로운 확률 사상을 만들어 내는 방법은 외적이며 확률 사상의 외적인 컨볼루션을 적용할 수 있다.

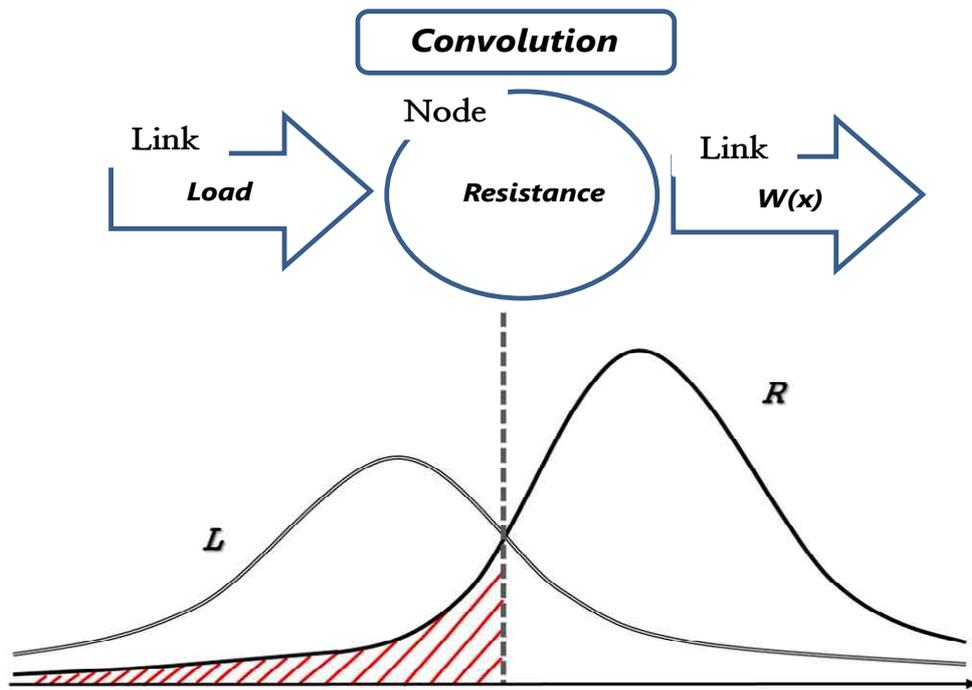


Fig. 4.1 Convolution in network that calculate link and node

네트워크에서  $L$ 은 절점으로 들어오는 연결선이 가지는 외력이고  $R$ 은

절점이 가지는 저항이다. 그리고  $W$ 는 절점에서 나가는 연결선이 가지는 저항을 거친 외력을 의미한다.

$$\begin{aligned}
 \int_{-\infty}^{\infty} W(x)dx &\Leftrightarrow \int_{-\infty}^{\infty} L(x) \times R(x)dx \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^x L(\tau) d\tau R(x)dx \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^0 L(\tau+t) R(\tau) dt dx \\
 &= \int_{-\infty}^0 L(t) \otimes R(-t) dt
 \end{aligned} \tag{2}$$

$L$ : Load in link  
 $R$ : Resistance in node  
 $W$ : Weight in link

$L, R$ 의 확률분포를 정규분포로 가정할 때 평균값을 대푯값으로 이용하면  $W$ 의 평균값은  $L-R$ 과 같고 컨볼루션 해법은 식 (2)와 같다. 이는  $W(x)$ 가 0보다 큰 영역을 컨볼루션을 이용해서 구한 식으로 이를 이용하면 다음 절점에서의 연결성을 확률적 범위로 산정할 수 있다.

## 4.2.2 균집화 모델 개발

### 1) 적용 범위

네트워크 구조의 균집화를 위해서는 각 절점들의 관계 정보인 연결선과 절점의 특성을 포괄적으로 고려하여 절점들을 대표할 수 있는 확률분포를 구해야한다. 방법론을 적용하기 위해 본 연구에서 사용된 네트워크의 범위는 다음과 같다.

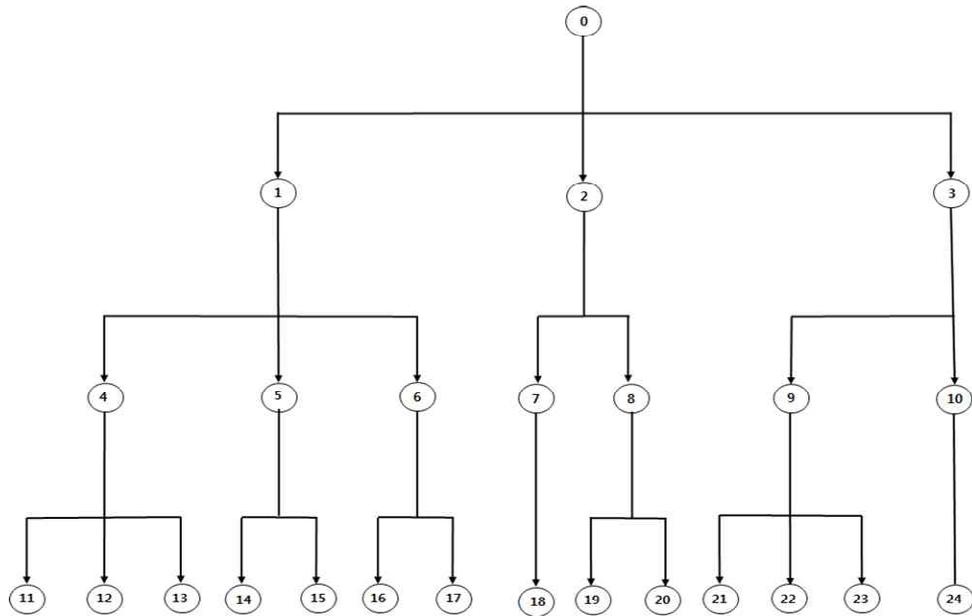


Fig. 4.2 Research scope of directed, hierarchy hub network

1. 본 연구에서 적용할 네트워크의 형태는 흐름을 가지는 일방향성을 가지는 네트워크 구조이다.
2. 절점은 특징적인  $R$ (저항)의 확률분포를 가진다.
3. 연결선은  $L$ (하중)의 확률분포를 가지며 방향성을 가진다.
4.  $R$ 과  $L$ 은 서로 대응하며 컨볼루션 연산 후에 연결선은  $W$ 의 확률분포를 가진다.

Fig. 4.2는 본 연구에서 제안한 모델에서 적용 가능한 범위에 있는 네트워크 구조이다. 최상위 0번 절점에서 흐름이 시작되며 각 절점은 흐름에 대한 저항을 가지고 있다. 최상위 0번 절점에서 물질의 흐름이 시작된다면 확률 재규격화를 이용한 군집화(Clustering with Probabilistic Convolution, CPC)를 이용해서 각 절점이 가지는 저항에 따라서 그 흐름을 아래 절점으로 전달시켜주는 것이 가능한지 여부를 각 절점별로 확

를 값을 표현할 수 있고 그 값에 따라서 최상위 절점과의 연결 정도를 정량화할 수 있다.

## 2) CPC 모델의 개발

Fig. 4.3과 같이  $i, j, k$  절점이 순서대로 방향성을 가지고 이어져 있는 경우  $i$  절점에서  $j$  절점에 전달하는 외력  $L$ 은  $j$  절점의 저항력  $R$ 과 연산되어  $k$  절점으로 전달되는  $W$ 가 산정된다.

$W$ 는  $k$  절점에서 다른 절점으로의 연결을 연산할 때  $L$ 로 치환될 수 있으며 해석은 단계별로 축차적으로 해석된다.

$$W = \max(L - R, 0) \quad (3)$$

식 (3)와 같이 연결선이 가지는 값은 절점에 작용하는  $L$ 에서  $R$ 을 뺀 값이며 음의 값이면 연결선이 발현되지 않은 것을 의미한다. 이때 임의의 하중,  $x$ 에 따라 절점에 작용하는 외력의 확률을  $L(x)$ , 저항의 확률을  $R(x)$ , 저항을 거친 외력을  $W(x)$ 라 하면,

$$\begin{aligned} W(x) &= W(L - R) \\ &= \left( \int_{-\infty}^x R(\tau) d\tau \right) \times L(x) \end{aligned} \quad (4)$$

식 (4)와 같다.  $W(x)$ 는 하중  $x$ 가 나타날 수 있는 모든 경우를 합해야 한다.  $L(x)$ 와  $R(x)$ 가 가질 수 있는 모든 확률적 경우를 합산할 경우  $W(x)$ 가 된다. 따라서 이를 전체적인 확률분포로 표현하기 위해 임의의 하중을  $-\infty$ 에서  $+\infty$ 까지 정리하면 연결선의 확률밀도함수는 다음과 같다.

$$\begin{aligned}
PDF_{W(x)} &= conv(PDF_{L(x)}, PDF_{R(x)}) \\
\int_{-\infty}^{\infty} L(x) \times R(x) dx &= \int_{-\infty}^{\infty} \int_{-\infty}^x L(\tau) d\tau R(x) dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^0 L(\tau+t) R(\tau) dt dx \\
&= \int_{-\infty}^0 L(t) \otimes R(-t) dt \tag{5}
\end{aligned}$$

임의의 절점에서의 네트워크의 임의의 절점이 가지는 수식으로 확장하면 식 (6)과 같다.

$$PDF_{W(x_n)} = conv(PDF_{L(x_{*t})}, PDF_{R(x_n)}) \tag{6}$$

$x_n$  : node n

$x_{start}$  : start node lined with node n

$x_{start}$  는 임의의 절점과 연결되어있는 시작점을 의미한다. 식 (6)의 연산이 시작절점에서부터 단계별로 이루어지기 때문에 전체 네트워크 구조에서 이루어지는 연산 과정은 식 (7)과 같다.

$$PDF_{W(x_k)} = conv(conv(PDF_{L(x_i)}, PDF_{R(x_j)}), PDF_{R(x_k)}) \tag{7}$$

$PDF$  : Probability Density Function

$conv(a,b)$  : convolution a and b

Fig. 4.3의 k 절점은 i 절점에서 시작된 흐름이 j 절점에서의 저항과 k 절점의 저항을 거치기 때문에 k 절점의 W 분포의 평균값은 j 절점의 W 분포의 평균값과 비교할 때 더 작다. 그리고 발현의 기준이 되는 W값이 0보다 작을 확률이 j 절점보다 크다.

Fig. 4.4는 시작 절점에서 컨볼루션 연산의 반복된 수행이 어떻게 진행되는지에 대한 순서도를 나타낸 것이다.

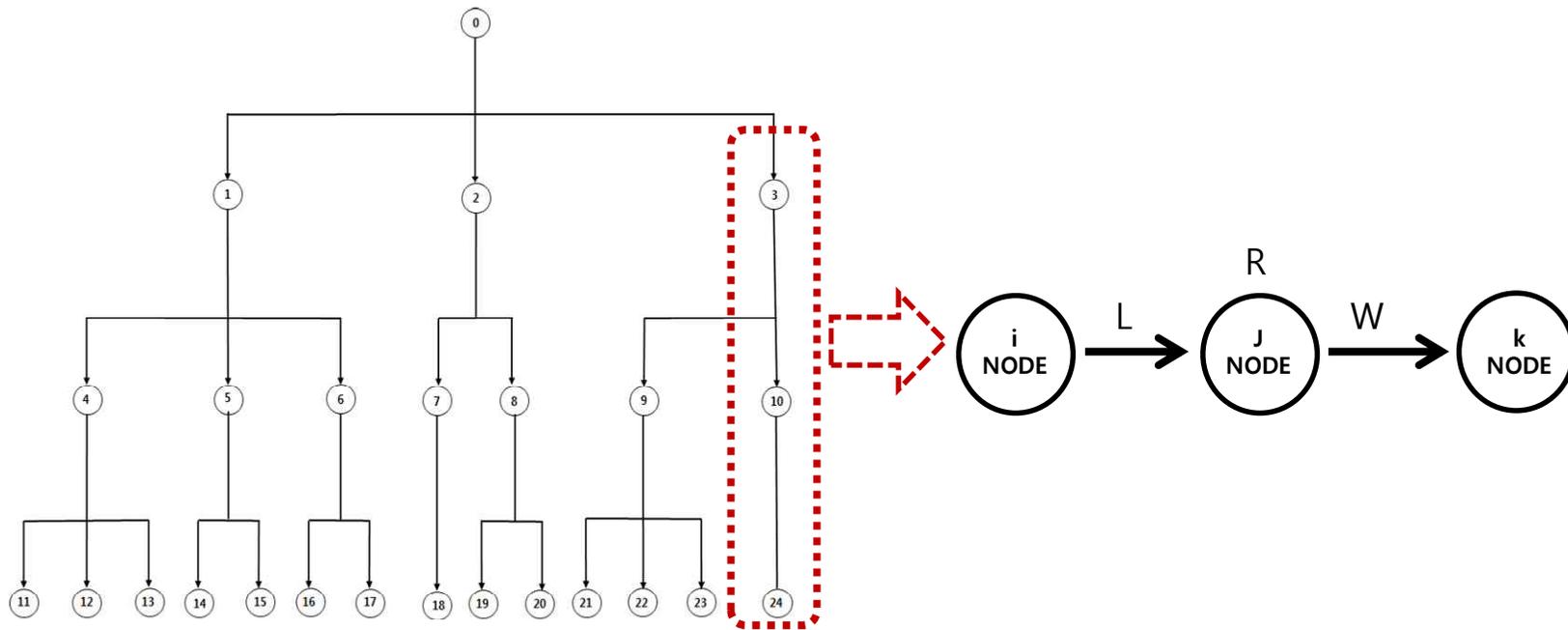


Fig. 4.3 Convolution in directed, hierarchy hub network

복잡한 방향성을 가진 네트워크를 축차적으로 컨볼루션을 수행함에 따라 각 절점이 가지는 연결선에 대한 확률적 분포를 얻을 수 있다.

각 절점마다 저항의 차이가 있겠지만 최상위 절점에서 하위 절점으로 단계가 내려갈수록 발현되지 않을 확률이 점점 커지게 된다. Fig. 4.4의 순서도에 따라 연산이 수행되면 절점이 가지는 연결선이 발현될 확률,  $W_{positive}$  을 각각 산정할 수 있고 이를 이용해서 최상위 절점과 밀접한 연관이 있는 절점들의 집단을 군집화할 수 있다.

$$W_{positive} = \int_0^{\infty} PDF_{node} \quad (8)$$

절점의  $W$ 가 0보다 큰 확률을 식 (8)과 같이 쓸 수 있으며, 이는 연결선의 발현 기준이 된다.

$$Group_{node_x} = Clustering(node_1, node_2, \dots, node_n, W_{positive} > C.C.) \quad (9)$$

$C.C.$ : Criteria of Clustering

식 (9)와 같이 군집화의 기준(Clustering of Criteria,  $C.C.$ )이 마련되면 1부터  $n$ 까지 절점의  $W_{positive}$  를 분석하여  $x$  절점과 가장 밀접하게 연관된, 그리고 특성이 유사한 절점들을 구분할 수 있다. 군집화의 기준은 절점들의 정보와 적용하고자 하는 문제에 따라 유연하게 설정할 수 있으며 그 기준에 따라 적합한 군집화 결과를 얻을 수 있다.

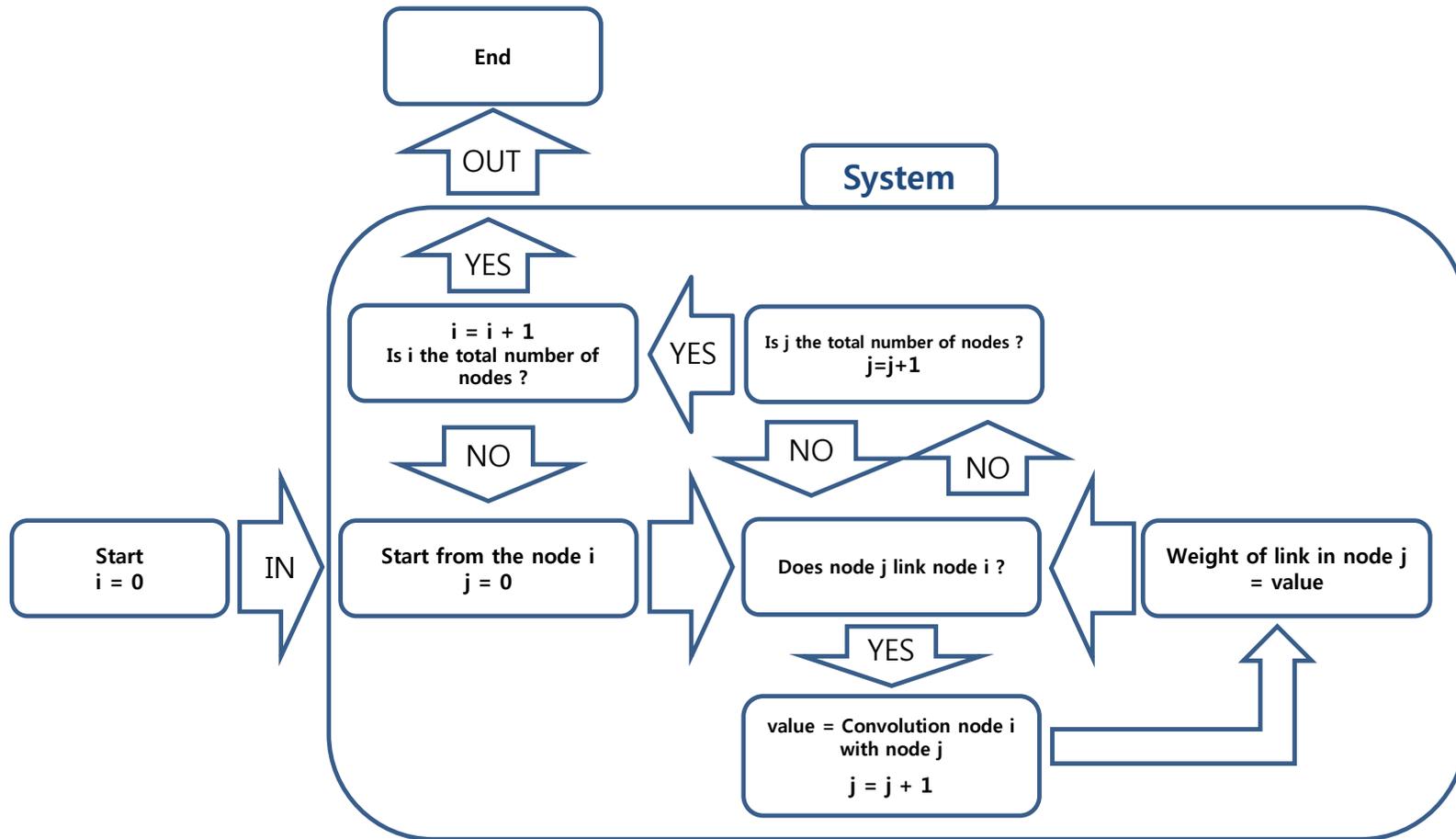


Fig. 4.4 System flow of Clustering with Probabilistic Convolution model

## 4.3 균집화 모델의 구현 및 고찰

### 4.3.1 CPC 모델의 구현

임의의 네트워크 구조를 가정하여 컨볼루션을 이용한 균집화 방법을 적용하고 그 결과를 난수를 이용한 수치해법인 몬테카를로 방법과 비교, 검증하며 다른 균집화 방법과의 차이를 분석해보고자 한다. Fig. 4.5의 네트워크는 방향성을 가지는 절점 11개와 연결선 10개로 구성된 네트워크이다. 0번 절점에서 응력의 흐름이 시작되는 구조 네트워크라고 가정을 한다.

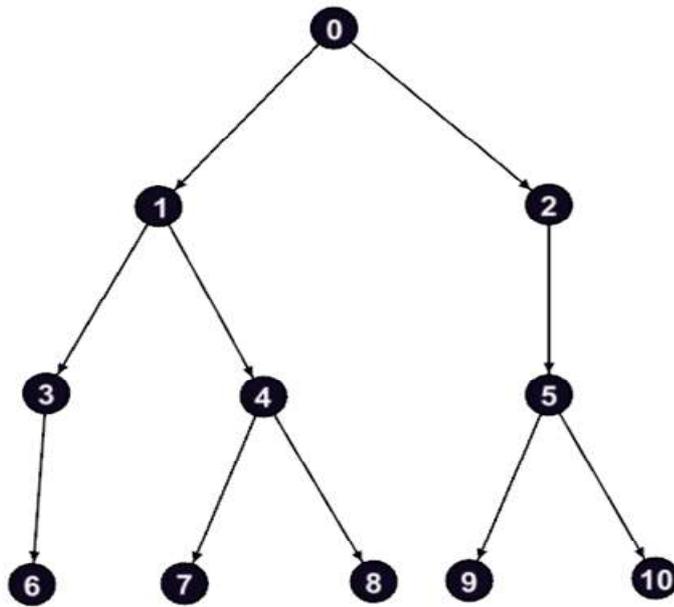


Fig. 4.5 Example of the network for CPC model

최상위 0번 절점에서 응력의 흐름이 시작되는 구조라면 분기점이 되는 0번 절점에서 1, 2번 절점, 1번 절점에서 3, 4번 절점, 4번 절점에서 7, 8번 절점, 5번 절점에서 9, 10번 절점으로 각각 갈라지는데 그 비는 각각 Table 4.1과 같다.

Table 4.1 Distribution of the node in the example of the network

start node	0		1		4		5	
end node	1	2	3	4	7	8	9	10
portion	0.7	0.3	0.6	0.4	0.7	0.3	0.4	0.6

11개 절점의 저항력  $R$ 의 확률 분포는 Table 4.2와 같고 응력을 전달하는 전달력  $L$ 의 확률 분포는 Table 4.1에서 가정한 비에 따르며 최초 0번에서 주어지는 연결선의 전달력은 10으로 하고 표준편차는 1로 하고 모든 분포는 정규분포(Normal Distribution)의 형태를 가진다고 가정하였다.

Table 4.2 Distribution of the Resistance in node

Node	Mean	Std.
0	2	1
1	2	$\sqrt{3}$
2	3	1
3	2	1
4	3	$\sqrt{3}$
5	2	1
6	3	$\sqrt{6}$
7	2	$\sqrt{3}$
8	3	1
9	2	$\sqrt{5}$
10	2	$\sqrt{2}$

가정된 자료를 이용하여 계산하면 1번 절점에 응력이 전달될 확률은 다음과 같다.

$$PDF_{W_1(x)} = conv\left(\frac{7}{10}PDF_{L_0(x)}, PDF_{R_1(x)}\right) \quad (10)$$

이때  $L_0(x)$  분포는 평균이 7(0번에서 주어지는 하중 10에 0.7을 곱한 값), 표준편차는 1인 정규분포이다. 식 (10)에서 구한 값을 이용한 3번 절점에 응력이 전달될 확률은

$$\begin{aligned} PDF_{W_3(x)} &= conv\left(\frac{6}{10}conv\left(\frac{7}{10}PDF_{L_0(x)}, PDF_{R_1(x)}\right), PDF_{R_3(x)}\right) \\ &= conv\left(\frac{6}{10}PDF_{L_1(x)}, PDF_{R_3(x)}\right) \end{aligned} \quad (11)$$

식 (11)와 같다. 같은 방법으로 각 절점에 응력이 전달될 확률을 구하는 방법은 Table 4.3과 같고 단계별 계산 방법을 Fig. 4.6과 같이 도식화 하였다.

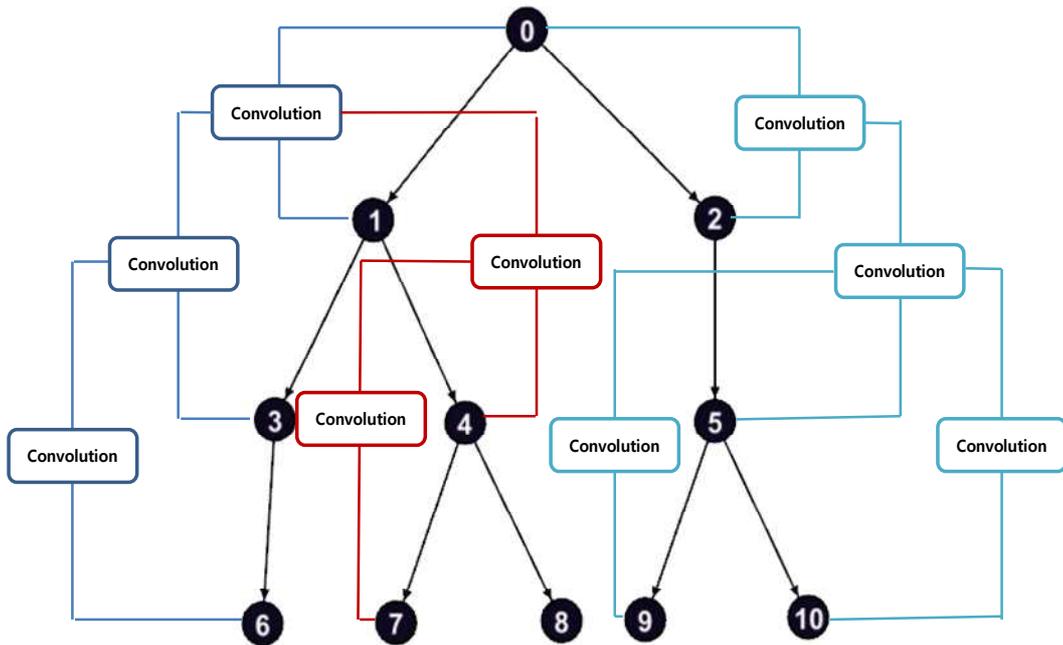


Fig. 4.6 Analysis process of the example

Table 4.3 Explanation to the probability of stress flow in node

Node	explanation
1	$PDF_{W_1(x)} = conv(\frac{7}{10} PDF_{L_0(x)}, PDF_{R_1(x)})$
2	$PDF_{W_2(x)} = conv(\frac{3}{10} PDF_{L_0(x)}, PDF_{R_2(x)})$
3	$PDF_{W_3(x)} = conv(\frac{6}{10} PDF_{L_1(x)}, PDF_{R_3(x)})$
4	$PDF_{W_4(x)} = conv(\frac{4}{10} PDF_{L_1(x)}, PDF_{R_4(x)})$
5	$PDF_{W_5(x)} = conv(PDF_{L_2(x)}, PDF_{R_5(x)})$
6	$PDF_{W_6(x)} = conv(PDF_{L_3(x)}, PDF_{R_6(x)})$
7	$PDF_{W_7(x)} = conv(\frac{7}{10} PDF_{L_4(x)}, PDF_{R_7(x)})$
8	$PDF_{W_8(x)} = conv(\frac{3}{10} PDF_{L_4(x)}, PDF_{R_8(x)})$
9	$PDF_{W_9(x)} = conv(\frac{4}{10} PDF_{L_5(x)}, PDF_{R_9(x)})$
10	$PDF_{W_{10}(x)} = conv(\frac{6}{10} PDF_{L_5(x)}, PDF_{R_{10}(x)})$

식(11)에 따르면 절점 0에 영향을 받는 절점 1이 응력의 영향을 받을 확률( $W_{positive}$ )은

$$\begin{aligned} W_{positive} &= \int_0^{\infty} PDF_{w_1} \\ &= 99.76\% \end{aligned} \quad (12)$$

식 (12)가 된다. 식 (12)의 결과는 0번 절점에서 응력의 흐름이 시작되면 1번 절점이 영향을 받을 확률이 99% 이상이라는 것을 의미하며 0번 절점과 아주 밀접한 연관을 가지고 있는 절점이라는 것을 의미한다. 앞선 수식과 Fig. 4.6의 프로세스를 통하여 산정된 각 절점에 응력이 전달될 확률은 Table 4.4와 같다.

Table 4.4 Probability of stress flow in node

Node	Probability of stress flow(%)
1	99.76
2	92.14
3	71.92
4	42.21
5	50.12
6	41.42
7	21.16
8	10.12
9	15.32
10	22.36

기준이 완화될수록 0번 절점과 같은 집단에 속하는 절점들이 많아진다. 1번, 2번 절점은 0번 절점의 영향 아래 90% 이상 영향을 받기 때문에  $W_{Negative}$ 가 90%보다 낮으면 항상 같은 집단으로 포함되며 4번, 6번 절점은 40%를 조금 상회하는 발현 확률 값에 따라 40% 이하가 되기 전까지는 포함되지 않는다. 본 방법을 통해서 적용하고자 할 때 관련된 문제 다수의 자료 고찰을 통해서 문제에 대한 해법을 찾기 위한 적절한 군집화의 기준을 설정할 수 있을 것으로 판단된다.

$$\begin{aligned}
 W_{-ative} < 90\%, \text{ Group} &= [\text{node}_0, \text{node}_1, \text{node}_2] \\
 W_{-ative} < 50\%, \text{ Group} &= [\text{node}_0, \text{node}_1, \text{node}_2, \text{node}_3, \text{node}_5] \\
 W_{-ative} < 40\%, \text{ Group} &= [\text{node}_0, \text{node}_1, \text{node}_2, \text{node}_3, \text{node}_4, \text{node}_5, \text{node}_6] \quad (14)
 \end{aligned}$$

본 방법을 통해서 적용하고자 할 때 관련된 문제 다수의 자료 고찰을 통해서 해법을 찾기 위한 적절한 군집화의 기준을 설정할 수 있을 것으로 판단된다. 절점과 연결선으로 모델링된 구조 문제에 적용된다면 응력의 흐름의 양상을 분석할 수 있을 것으로 기대된다.

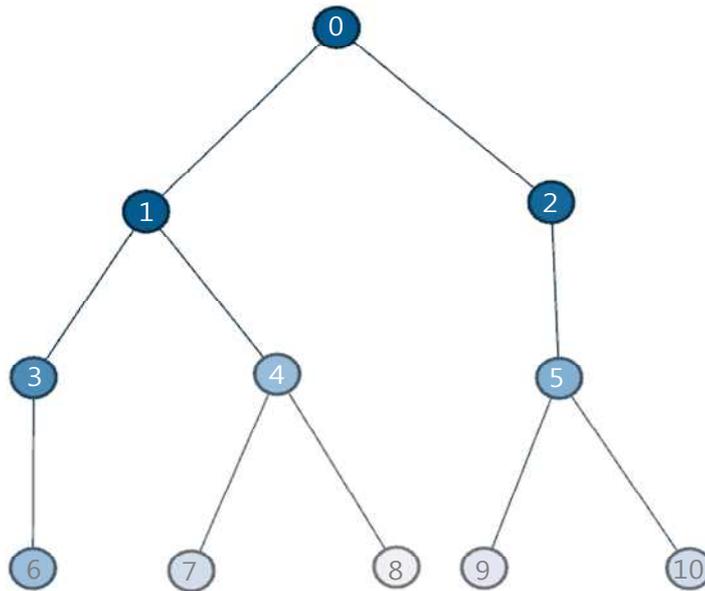


Fig. 4.7 Result of the stress flow using CPC model

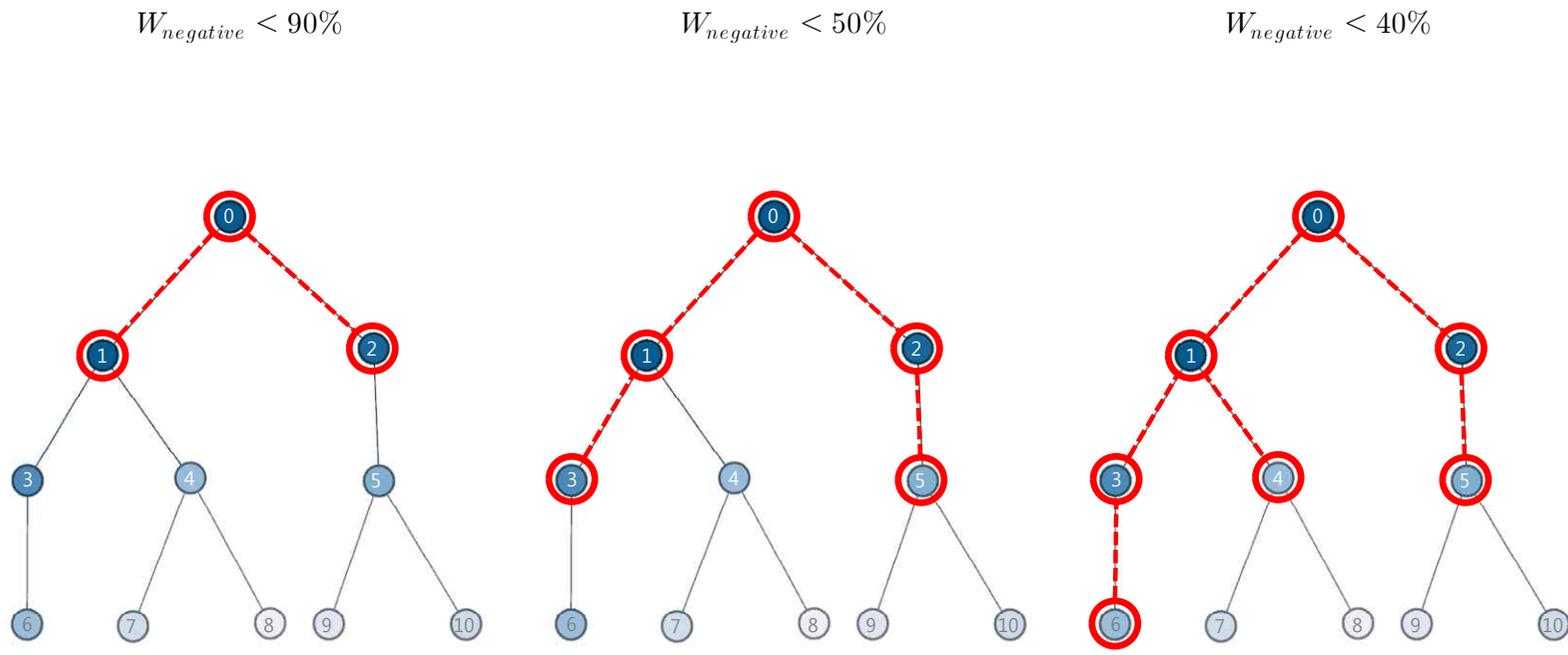


Fig. 4.8 Result of the stress flow using CPC model with the criteria of clustering

### 4.3.2 CPC 모델의 검증

결정론적 모델과 비교할 때 변수의 범위가 확률적 분포를 가지는 확률 모델에서는 분석적 해를 구하는 것이 어렵다. 본 연구에서 제안한 CPC 모델 역시 각 절점의 연결선 발현에 대한 값의 확률적 분포만을 제공하기 때문에 분석적인 해를 찾는 것이 불가능하다. 따라서 분포가 주어진 난수를 반복적으로 발생시켜 수치적으로 모의하여 답을 찾는 몬테카를로 시뮬레이션(Monte-Carlo Simulation)을 통해서 CPC 모델의 결과를 비교 검증할 수 있다. 일반적으로 분석적 해를 통한 검증이 어려울 경우 난수에 가까운 몬테카를로 시뮬레이션의 수치적 해법을 통한 검증은 신뢰할 수 있는 방법이다 (JM Burt et al., 1971; Romualdo Pastor-Satorras and Alessandro Vespignani, 2000). 본 연구에서 수행한 몬테카를로 시뮬레이션을 통해서 발생한 난수는 1,000,000회 계산이었으며 컨볼루션을 이용한 각 절점의 연결선 발현확률인 Table 4.4의 결과와 비교하였다. 발현확률을 구하는 방법은 다음 식과 같다.

$$W_{positive} = \frac{n_p}{1,000,000} \quad (15)$$

$n_p$  : the number of generating nodes

여기서,  $n_p$ 는 난수를 통한 해석 시 절점의 연결선이 발현 되는 횟수를 의미하고 총 1,000,000번의 시행가운데 그 횟수를 통해서 각 절점의 발현 확률을 Table 4.5와 같이 구할 수 있다.

Table 4.5 Comparison with Monte-Carlo simulation

Node	Probability of expression for stress flow		
	CPC <sup>1</sup>	MCS <sup>2</sup>	error
1	0.99761	0.99389	0.003743
2	0.92143	0.91935	0.002262
3	0.71922	0.71711	0.002942
4	0.42211	0.42116	0.002256
5	0.50125	0.49709	0.008369
6	0.41424	0.41382	0.001015
7	0.21164	0.20956	0.009926
8	0.10121	0.10038	0.008269
9	0.15322	0.15266	0.003668
10	0.22363	0.22148	0.009707

1) CPC : Clustering by Probabilistic Convolution

2) MCS : Monte-Carlo Simulation

본 연구에서 개발한 컨볼루션을 이용한 군집화 방법(Clustering by Probabilistic Convolution)과 몬테카를로 시뮬레이션을 이용한 결과를 유효숫자 5자리까지 비교하였으며 몬테카를로 시뮬레이션의 상대오차의 방법에 따라 오차를 구하면 다음과 같다 (LJ Ringer, 1971; JM Burt et al., 1971).

$$\begin{aligned}
error &= \sqrt{\sum_{k=1}^{10} \left( \frac{CPC_k - MCS_k}{MCS_k} \right)^2} \times 100 \\
&= 0.5216\%
\end{aligned}
\tag{16}$$

식 (16)을 이용해 Table 4.5의 10개 절점의 연결선 발현 확률을 각각 비교하고 합사하면 99.4%이상의 정확도를 보여준다. 11개의 절점으로 구축된 간단한 구조를 대상으로 하였고 분포가 고정되어 있기 때문에 몬테 카를로의 정확도가 높게 나오긴 했지만 복잡한 구조에 절점이 다양한 분포를 가진다하더라도 정확도는 크게 떨어지지 않을 것이며 확장해서 사용할 수 있을 것으로 기대된다.

### 4.3.3 기존 모델과 비교

본 연구에서 소개된 방법을 기존의 가중 허브 네트워크 군집화 연구들과 비교해보면 Table 4.6과 같다. Fig. 4.5의 예제 네트워크에서 최상위 0번 절점과 연결되어있는 연결선의 합을 10으로 하고 Table 1.1의 비율과 같이 연결선의  $L$ 을 가정하였고 Table 2의 각 절점의  $R$  평균값을 이용하여  $L-R$ 로 최종적으로 산정된 연결선의 값이 적용된 결정론적 가중 네트워크를 기준으로 분석하였다.

Table 4.6 Comparison with the existing models

method	CPC (50%)	Clauset- Newman-Moore	Wakita-Tsurumi	Girvan-Newman
Group	0, 1, 2, 3, 5	0, 1, 2, 3, 5	0, 1, 2, 3	0, 1, 2, 3, 5

Clauset-Newman-Moore 모델과 Girvan-Newman 모델은 CPC와 동일한 절점을 군집하였다. Wakita-Tsurumi의 모델은 5번 절점을 제외하였는데 5번 절점은 CPC를 이용한 방법에서도 1, 2, 3, 5번 절점 가운데 0번 절점과 가장 연관도가 낮은 절점이다. 따라서 군집화의 기준을 5번 절점의 응력의 영향을 받을 확률인 50.12%보다 높게 설정하면 Wakita-Tsurumi의 모델과 같은 군집 결과를 얻을 수 있다. 5번 절점과 5개의 절점 가운데 2번째로 낮은 3번 절점의 응력의 영향을 받을 확률 차이는 약 21% 이상으로 나머지 4개 절점과 같은 집단으로 구분하기에는 범위가 크며 5번째로 낮은 4번 절점과의 차이가 약 7.9%로 오히려 더 높다.

기존 모델은 소집단을 추출하기 위해 전체 네트워크의 부하량을 산정해서 절점의 응력의 영향을 받을 확률을 구할 때마다 반복적으로 산정해야하기 때문에 절점수가 많은 네트워크에서는 효율적인 방법이 아닐 수도 있다 (R Guimera et al., 2003; Peter J. Mucha et al., 2010). 그러나 CPC의 경우 축차적으로 절점의 응력의 영향을 받을 확률을 구하며 전체 네트워크 구조와 관계없이 연계된 절점간의 확률분포만 합산하는 분산처리가 가능하기 때문에 전체 연산의 효율도를 높일 수 있다는 장점을 가진다.

$$\begin{aligned} \text{number of iteration in CPC} &= \text{number of links} \\ \text{number of iteration in the existing method} &= {}_n C_2 = \frac{n(n-1)}{2} \quad (17) \end{aligned}$$

$n$ : the number of nodes

식 (17)은 각 모델들의 연산 반복 횟수를 나타낸 것이다. CPC 모델에서 연산 반복 횟수는 절점간의 분산처리에 따라 연결선의 수만큼 시행하면 되지만 기존 모델의 경우 각각의 절점간의 연결 관계를 모두 산정해야하기 때문에 절점 수가 많아질수록 반복 횟수가 기하급수적으로 증가한다.

CPC 방법과 기존의 방법을 동일한 구조에서 단계별로 비례적으로 절점 수만 늘려가면서 반복 횟수를 비교해보면 Fig. 4.9와 같다. 절점 수가 많아질수록 연산 반복 횟수의 차이가 점점 커지는 것을 확인할 수 있다.

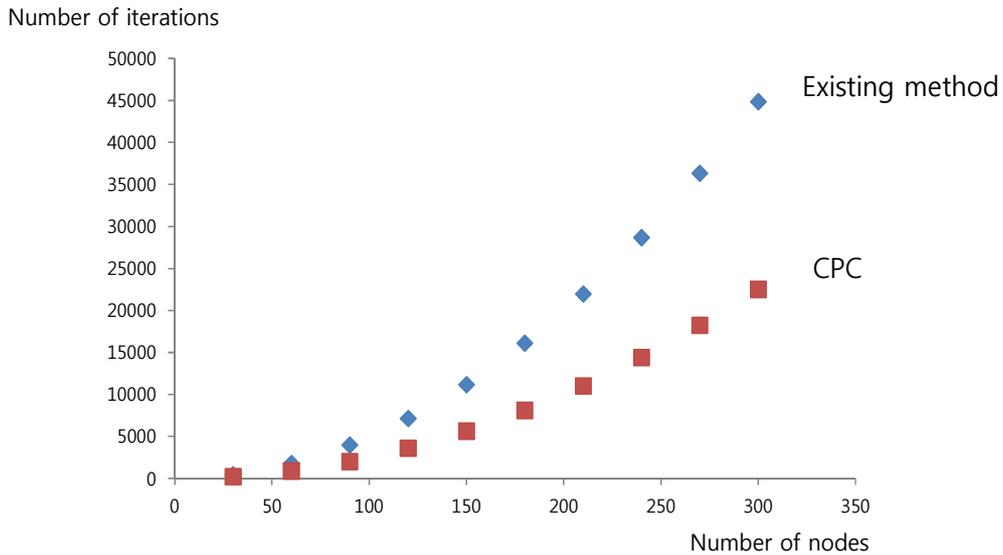


Fig. 4.9 Comparison of Number of iteration with the existing method

기존의 방법과 비교할 때 CPC 방법은 경계조건에 따라 절점과 연결선의 확률론적 값을 모두 반영할 수 있고 기준에 따라서 정확한 균집 결과를 산정할 수 있을 뿐만 아니라 연산의 효율성에 있어서도 장점을 가진다.

#### 4.4 소 결

확률론적 범위를 가지는 네트워크는 결정론적 값을 가지는 네트워크와 비교할 때 결정론적 해답을 도출하기는 어렵지만 확률적 연산이 수행

된다면 실측값에 대한 정확한 범위와 확률분포를 제시할 수 있다.

본 연구에서는 허브 네트워크의 절점과 연결선에 저항, 외력과 같은 변수를 적용한 모델을 만들었다. 그리고 저항, 하중의 확률분포를 가지는 절점 및 연결선의 정보에 대한 컨볼루션 연산을 통하여 새로운 연결선의 정보를 확률분포로 규격화하였다. 이 모델을 적용하여 방향성을 가지는 임의의 허브 네트워크 구조를 대상으로 확률분포를 분석하여 그 절점의 연결선 발현 확률을 해석하고 유사한 절점들을 군집화 하는데 이용하였다. 이 방법은 신뢰할 수 있는 수치적 해법 가운데 하나인 분포가 동일한 난수를 통한 몬테카를로 시뮬레이션 방법과 비교했을 때 방향성을 가지며 절점과 연결선이 확률을 가지는 제한된 네트워크에 적용한 방법론이지만 99.4% 이상의 정확도를 확보할 수 있었다. 그리고 기존의 다른 결정론적 네트워크의 군집화 방법과 비교했을 때 군집화의 기준에 따라 유연한 결과를 얻을 수 있으며 정확한 군집화가 가능함을 보였다.

확률론적 네트워크에서 이종의 확률적 연산을 통한 새로운 확률의 규격화는 경계조건이 고정된 네트워크 내의 요소들을 해석하기 위한 중요한 도구가 될 수 있음을 확인하였다. 방향성을 가지는 불균등한 구조를 가진 허브 네트워크의 내부 변화를 고찰하고자 할 때 절점 간의 확률적 군집화 방법론은 다양한 사례에 적용할 수 있을 것으로 기대된다.

## 참고문헌

1. Burt, J.M. and Garman M.B., 1971. Conditional Monte Carlo: A simulation technique for stochastic network analysis. *Management Science* 18(3): 207-217.
2. Dechter, R., and J. Pearl, 1989. Tree clustering for constraint networks. *Artificial Intelligence* 38(3): 353-366.
3. Fefferman, C., 1970. Inequalities for strongly singular convolution operators. *Acta Mathematica* 124(1): 9-36.
4. Friedman, N.. 2004. Inferring Cellular Networks Using Probabilistic Graphical Models. *Science* 303.5659: 799-805.
5. Friedman, N., D. Geiger, and M. Goldszmidt, 1997. Bayesian Network Classifiers. *Machine learning* 29(2-3): 131-163.
6. Girvan, M. and M.E.J Newman, 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99(12): 7821-7826.
7. Gudkov, V., J.E. Johnson, and S. Nussinov, 2002. Graph equivalence and characterization via a continuous evolution of a physical analog. arXiv preprint cond-mat/0209112.
8. Leskovec, J., K.J. Lang, and A. Dasgupta, 2007. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics* 6(1): 29-123.
9. Leskovec, J., K.J. Lang, and M. Mahoney, 2010. Empirical comparison of algorithms for network community detection. *Proceedings of the 19th international conference on World wide web*, 631-640. New York, USA.
10. Marcos, G.Q., L. Zhao, L. Ronaldo, and A.F. Roseli, 2008. Particle

- competition for complex network community detection. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 18(3): 033107–033107.
11. Newman, M.E.J. and M. Girvan, 2004. Finding and evaluating community structure in networks. *Physical review E*. 69(2): 026113.
  12. O’Neil, R., 1963. Convolution operators and  $L(p, q)$  spaces. *Duke Mathematical Journal* 30(1): 129–142.
  13. Peter, J. M., T. Richardson, K. Macon, A.P. Mason, and J. Onnela, 2010. Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. *Science* 328(5980): 876–878.
  14. Pothén, A., H.D. Simon, and K.P. Liou, 1990. Partitioning sparse matrices with eigenvectors of graphs. *Journal on Matrix Analysis and Applications* 11(3): 430 - 452.
  15. Reichardt, J. and S. Bornholdt, 2004. Detecting fuzzy community structures in complex networks with a Potts model. *Physical Review Letters* 93(21): 218701.
  16. Reichardt, J. and S. Bornholdt, 2006. Statistical mechanics of community detection. *Physical Review Letters* 74(1): 016110.
  17. Ringer, L.J., 1971. A statistical theory for PERT in which completion times of activities are inter-dependent. *Management Science* 17(11): 717–723.
  18. Romualdo, P. and V. Alessandro, 2001. Epidemic spreading in scale-free networks. *Statistical Mechanics* 86(14): 3200–3203.
  19. Roy, S., D. Saha, D. Bandyopadhyay, T. Ueda, and S. Tanaka, 2003. A network-aware MAC and routing protocol for effective load balancing in ad hoc wireless networks with directional antenna. *Proceedings of the 4th ACM international symposium on*

Mobile ad hoc networking & computing 88–97. New York, USA.

20. Wu, F. and B.A. Huberman, 2004. Finding communities in linear time: a physics approach. *The European Physical Journal B* 38(2): 331–338.
21. Zhou, H., 2003. Distance, dissimilarity index, and network community structure. *Physical Review E*. 67(6): 061901.



## 제 5장 허브 네트워크 추출과 분석 모델의 적용

본 장에서는 혼합 네트워크의 특성을 가지는 대상의 구조 신뢰성과 확산 문제를 해석하기 위하여 허브 네트워크를 추출하는 방법, 취약성 분석, 군집화 분석 등의 방법론으로 적용하고자 한다. 연속체 구조물과 뼈대 구조물을 네트워크로 치환하여 네트워크 추출과 군집화 분석의 방법론으로 각각 응력 흐름을 모의하고 파괴 기작을 해석하였다. 그리고 가속전염병의 확산 차단 방안을 네트워크 추출과 취약성 분석 등의 방법으로 분석하고 제안하였다.

### 5.1 연속체 구조물의 응력 흐름에 대한 군집화 적용

연속체 구조물을 요소와 연결선으로 치환하면 네트워크와 유사한 형태가 된다. 연결선을 통해서 응력이 전달되는 과정을 네트워크 해석 방법으로 모의할 수 있으며 외부 하중에 유사하게 반응하는 요소들을 군집화할 수 있다.

#### 5.1.1 연속체 구조물의 이산화

연속체에서는 물질 내의 작은 입자들이 균일하게 분포되어 있다고 가정한다. 그러나 연속체 해석의 경우 복잡한 형상을 가지거나 다양한 경계조건에 따라 포괄적 해석이 필요한 대상에 대해서 해석적 방법의 해를 구하기가 어렵다.

연속체 구조를 이산화하면 해석 속도의 효율성도 향상되며 복합적인

요인에 따른 해석 결과의 이해도를 높일 수 있다 (Lee, 2009). 구조물을 이산된 구조체로 해석하는 방법으로 트러스 모델 등이 있으나 대상 영역을 트러스 구조로 치환하는 전처리 과정에 있어서 설계자의 경험이나 직관 등의 주관이 개입되어 정형화된 결과를 얻어내기 힘들다. 그리고 내부 임의의 지점의 응력을 확인할 수 없으며 전체적인 거동만 해석할 수 있다는 단점을 가진다. 이를 해결하기 위한 방법으로 해석의 전처리 과정으로 대상 영역을 불규칙 삼각망으로 구성하거나 등가 트러스 모델 등을 이용해서 해석을 시도한 연구가 있다 (Lee, 2009; Lee, 2008; Yoon, 2001). 이 연구들은 이산된 구조체의 각 요소들의 위상 체계를 확립할 수 있었으며 복잡한 구조의 해석할 때 결과의 이해도나 해석의 효율성을 향상시키는 등의 기여를 하였다.

불규칙 삼각망이나 등가 구조 등으로 이산화된 구조물은 절점과 연결선으로 구성된 네트워크와 구조가 유사하다. 각 절점을 이어주는 연결선은 트러스 요소의 인장과 압축을 나타내고 그 값에 따라서 인접한 다른 절점과 연결선에 영향을 미친다. 인접한 절점들을 중심으로 선호적 연결을 하고 있지만 각 연결선이 값을 가지고 있기 때문에 균등한 무작위 네트워크 구조보다 불균등한 가중 허브 네트워크와 유사한 특성을 가진다.

본 장에서는 허브 네트워크의 추출, 군집화 방법론 등을 이산화된 연속체의 구조 문제에 적용하여 응력 흐름을 탐색하고자 한다. 먼저 트러스 요소로 이산화된 연속체에서 수행지수를 이용하여 가장 효율적인 최적구조를 탐색하는 방법과 허브 네트워크 추출 방법의 비교를 통하여 구조 문제에서 적용성을 검증하였다. 그리고 연속체를 네트워크 요소로 치환하고 각 요소들이 확률적 범위를 가질 때 3장의 군집화 방법론을 적용해서 응력 흐름을 분석하였다. 두 가지 방법을 통한 해석은 네트워크 내 연결선의 중요도를 정립할 수 있기 때문에 콘크리트 구조내의 철근 배근과 같은 실용적인 문제에도 충분히 활용가능 할 것으로 판단된다.

## 5.1.2 허브 네트워크 추출과 진화론적 최적화 구조 기준

### 1) 구조물의 트러스 모델링

다수 절점과 부재로 복잡하게 연결된 트러스의 해석을 위해서는 수치 계산을 이용한 행렬해법을 이용한다. 본 연구에서는 각 요소의 국소 좌표와 전체 좌표에 대한 관계를 통해서 강성 행렬을 구하고 식(1)과 같이 변위와 외력의 관계를 이용한 평형방정식을 이용하여 풀이하였다.

$$[f] = [k][u] \quad (1)$$

임의의 트러스 요소가 오일러 좌표계와 다른 방향으로 Fig. 1과 같이 기울어져 있을 때 국소좌표계( $x, y$ )와 전체좌표계( $X, Y$ )로 구분되고 두 절점은 외력과 변위를 가진다.

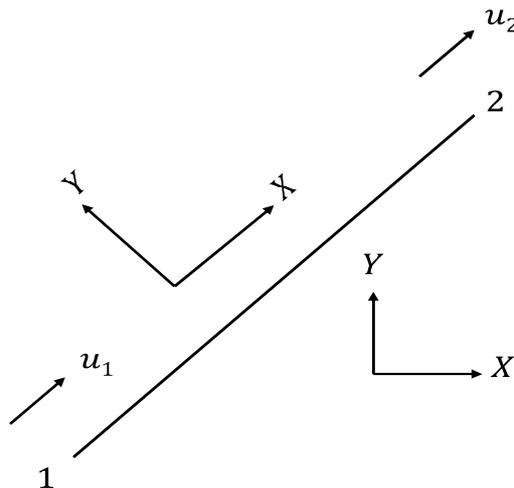


Fig. 5.1 Each truss element with coordinates

경계조건을 이용하고 2차원으로 확장하면 식(1)을  $4 \times 4$ 의 행렬식으로 정리할 수 있다.

$$\begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{bmatrix} = \frac{EA}{L} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \end{bmatrix} \quad (2)$$

그리고 국소 좌표계  $x, y$  축에 대한 방향코사인을  $l_1, m_1 ; l_2, m_2$  으로 하면 국소좌표계와 전체좌표계의 응력 관계는 식(3)과 같다.

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} l_1 & m_1 \\ l_2 & m_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3)$$

트러스에 작용하는 변형에너지는 식(4)과 같고

$$U = \frac{1}{2} (\overline{X_1 u_1} + \overline{X_2 u_2}) \quad (4)$$

위의 식들을 대입하면 다음과 같다.

$$U = \frac{EA}{2L} (u_1 v_1 u_2 v_2) \begin{bmatrix} l_1 & 0 \\ m_1 & 0 \\ 0 & l_2 \\ 0 & m_2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} l_1 & m_1 & 0 & 0 \\ 0 & 0 & l_2 & m_2 \end{bmatrix} \begin{pmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \end{pmatrix} \quad (5)$$

각 요소들의 강성행렬과 하중을 모두 합한다면 전체 구조의 강성행렬과 하중을 구할 수 있고

$$[k] = \begin{bmatrix} l^2 & lm \\ lm & m^2 \end{bmatrix} \quad (6)$$

변위를 이용하면 식(1)의 평형방정식에 대입해서 각 트러스 요소의 축력을 구할 수 있다.

$$S = [l, m] \begin{pmatrix} u_2 - u_1 \\ v_2 - v_1 \end{pmatrix} \quad (7)$$

축력은 인장과 압축이 있으며 값을 통해서 상대적인 위상을 비교할 수 있다.

## 2) 진화론적 구조 최적화 기준

절점과 수많은 요소들로 구성되는 트러스 구조는 주변 요소의 관계와 경계조건에 따라서 각기 다른 위상을 가진다. 응력 흐름에 대해서 해석을 하고 외력에 저항하지 못하는 요소를 단계별로 삭제하는 최적화 과정을 통해서 응력 흐름을 탐색한다. 그리고 이 방법론을 콘크리트 구조물에 적용할 때 철근 배근에 대해서 해석이 가능하다. 단, 단계별 불필요한 요소의 삭제시 필수적인 요소와 그렇지 않은 요소로 구분할 수 있는 최적의 기준이 명확하지 않다. 따라서 본 연구에서는 각 단계별 부피와 변형에너지를 이용해서 가장 효율적인 형상을 결정할 수 있는 수행지수 (Performance Index)를 이용하여 합리적인 종료기준을 이용한 최적화 과정을 수행하였다 (Hinton, 1995).

$$\frac{\sigma_{ele}}{\sigma_{max}} > RR_i \quad (8)$$

$RR$ 은 삭제율이며 식(8)을 통해서 이 기준에 부합하지 않는 요소는 불필요한 요소로 매 단계별로 삭제하였다. 그리고 수행지수는 식(9)와 같고,

$$PI = \frac{1/W_i V_i}{1/W_0 V_0} = \frac{W_0 V_0}{W_i V_i} \quad (9)$$

변형에너지의 밀도와 전체 요소의 수를 이용하여 단계별 모델의 효율성을 평가할 수 있다.

### 3) 최적 요소 추출

Fig. 5.2와 같이 분포하중을 받는 단순보를 불규칙삼각망으로 이산화하고 연결선을 트러스 요소로 치환하면 Fig. 5.3과 같이 412개의 절점과 718개의 연결선을 가진다.

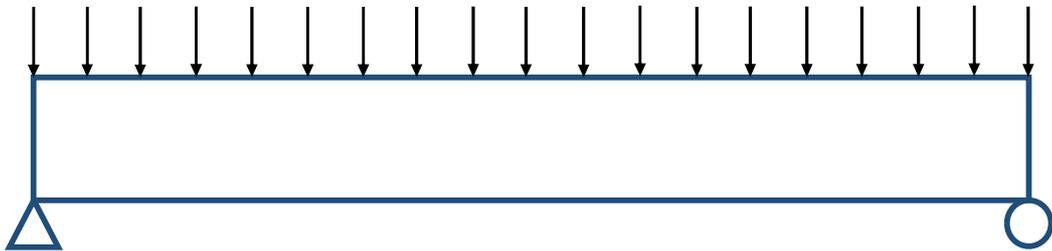


Fig. 5.2 Simple beam with distributed load

각 트러스 요소는 직접강도법으로 계산하면 인장력과 압축력을 가지게 되고 수행지수가 최대가 될 때는 전체 요소 가운데 30%의 요소만을 이용해서 기초 구조를 나타냈을 때 이다.

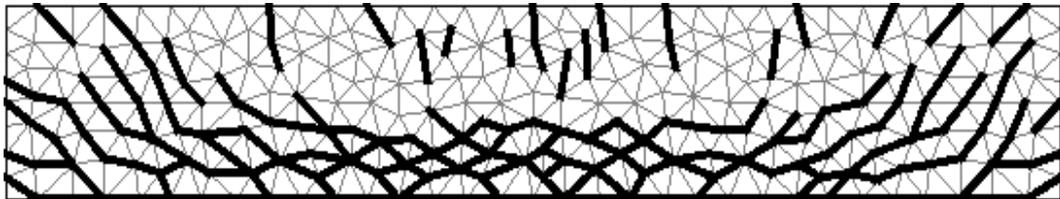


Fig. 5.3 Stress path in simple beam with maximum value of performance index

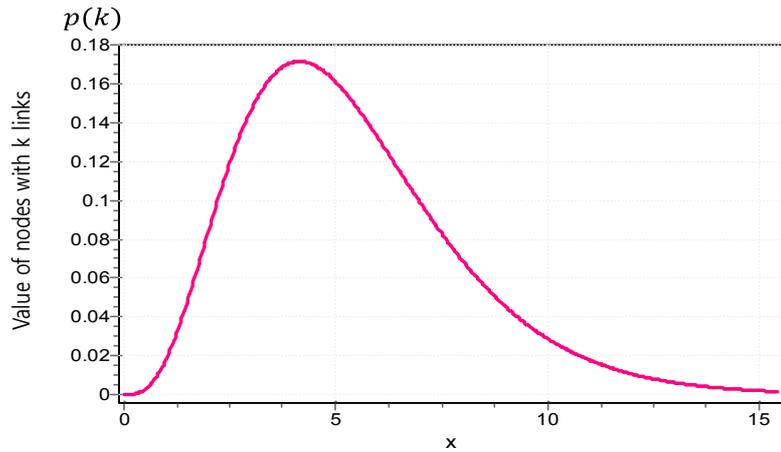


Fig. 5.4 Connectivity distribution in truss network

각 요소가 값을 가지고 있기 때문에 압축력은 모두 0으로 하면 값이 작은 불필요한 요소들은 무작위 네트워크와 같이 분포하고 값이 큰 요소들은 허브 네트워크와 같이 분포하게 된다. 불균등한 허브 네트워크라고 하면 연결선 수 분포를 감마분포로 시뮬레이션 할 수 있다. Fig. 5.4의 형상계수는 4.34, 규모계수는 1.24인데 허브 네트워크를 추출하면 임계치는 31.6%가 되며 Fig. 5.4에서와 같이 수행지수가 최대가 될 때의 제거율인 40.5%와 약 1.1%의 오차가 난다.

가중 네트워크의 연결선 값 분포를 이용하면 불필요한 요소는 제거하고 수행지수가 가장 높은 시기의 최적의 요소로 구성된 트러스 구조를 추출할 수 있다. 연속체를 네트워크로 구성하였을 때 값이 큰 요소들은 전체 구조에서 중요한 역할을 하고 값이 작은 요소들은 그렇지 않기 때문에 중요한 역할을 하는 요소들만 추출하여 가장 효율적인 구조를 구성할 수 있다.

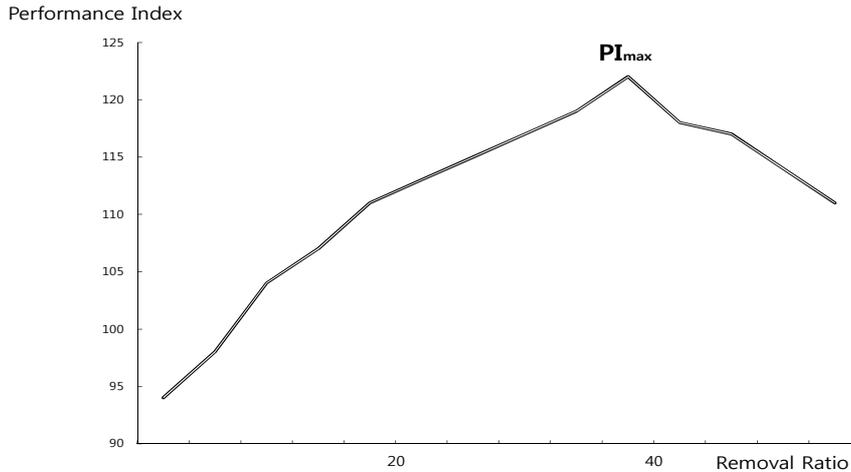


Fig. 5.5 Performance index with removal ratio

수행지수는 단계별로 구성된 모델들의 변형에너지와 부피 등을 고려해서 매번 산정해야하는 번거로움이 있지만 트러스 모델을 네트워크로 구축하고 연결선수 분포를 이용하면 임계치를 효율적으로 산정할 수 있다는 장점을 가진다.

### 5.1.3 응력 흐름의 균집화 분석

연속체를 절점과 연결선을 가진 네트워크로 치환하여 해석할 수 있다면 모델을 확장하여 각 요소가 확률적 요인을 가질 때 구조 내부의 응력을 해석할 수 있다. 외부 하중 조건에 따라서 연속체를 치환한 네트워크 모델은 연결선의 값이 서로 다른 불균등한 구조가 된다. 따라서 외부에서 응력의 흐름이 발생하면 지점의 반력과 반응하여 불균등한 방향으로 흐름이 진행된다. 응력 흐름에 따라 인장력의 발현 확률이 유사한 요소들을 균집화하면 연속체 내부의 응력 흐름을 탐색할 수 있다. 그리고 연속체 내부의 각 요소가 확률적 범위를 가질 때도 확장하여 사용할 수 있

다. 본 적용 사례에서는 연속체를 네트워크로 치환하고 각 요소가 확률적 범위를 가질 때 군집화 방법론을 적용하여 응력 흐름을 탐색하고자 한다. 복잡한 형태나 신뢰성 해석 등에서 구조 내의 속성을 효율적으로 분석할 수 있기 때문에 활용도가 높을 것으로 기대된다.

### 1) 응력 흐름 모델

군집화 분석을 위해서 지점의 반력과 각 외력에서 시작되는 응력의 흐름을 분석하기 위해 요소간의 기작을 정리할 필요가 있다. 먼저 트러스 네트워크의 임의의 절점으로 유입되는 응력의 합과 추출되는 응력의 합은 동일하다.

$$\sum F_{\in} = F_{out} \quad (10)$$

그리고 각 요소들은 국소좌표계를 따르기 때문에 주변 요소와의 연산을 위해서는 전체 좌표계의  $x, y$  축으로의 변환이 필요하다.  $x, y$ 축에 대한 방향 코사인을 각각  $l, m$ 이라 하면

$$\begin{aligned} F &= F_x + F_y \\ F_x &= Fl \\ F_y &= Fm \end{aligned} \quad (11)$$

임의의 부재의  $x$  방향 응력은 연결된 각 부재의  $x$ 축과의 방향 코사인의 비에 따라 응력이 흘러간다.

Fig. 5.6과 같이 부재  $i$ 의  $x$  방향의 응력은 해당 부재의 방향코사인에 그 부재와 연결된 모든 부재의 방향코사인의 합을 나눈 식(12)와 같다.

$$F_{i_x} = \frac{\cos \theta_i}{\sum_{k=1}^n \cos \theta_k} \times F_x \quad (12)$$

같은 방법으로  $y$  방향의 응력은 다음과 같다.

$$F_{i_y} = \frac{\sin \theta_i}{\sum_{k=1}^n \sin \theta_k} \times F_y \quad (13)$$

응력의 총합은 정해져 있기 때문에 각 부재의 방향의 비를 이용하면 각 부재로 흘러가는 응력의 합을 산정할 수 있다.

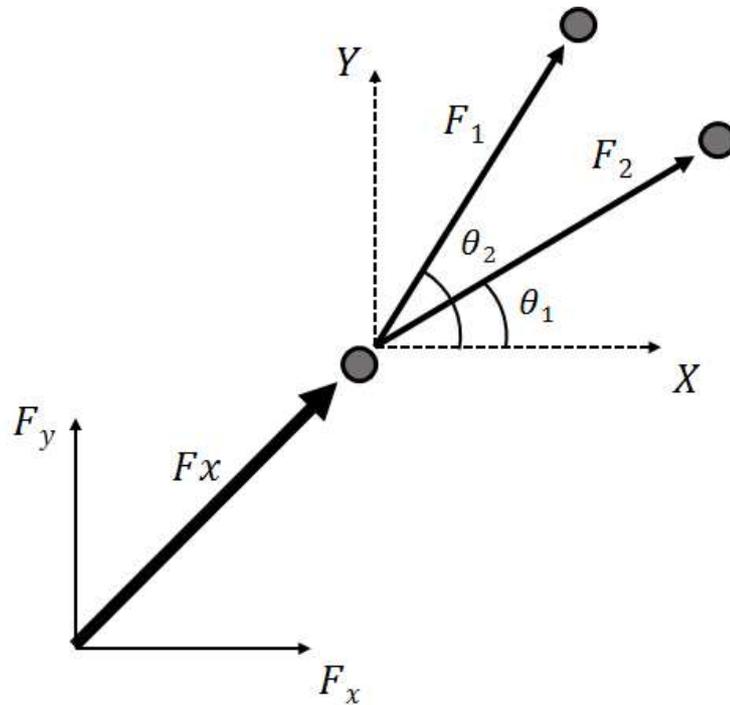


Fig. 5.6 Load of node in stress path model

## 2) 응력 흐름의 군집화

Fig. 5.7은 군집화 분석 방법을 단계적으로 나타낸 것이다. 확률적 응력 흐름의 군집화를 위한 모델 구성은 다음과 같다.

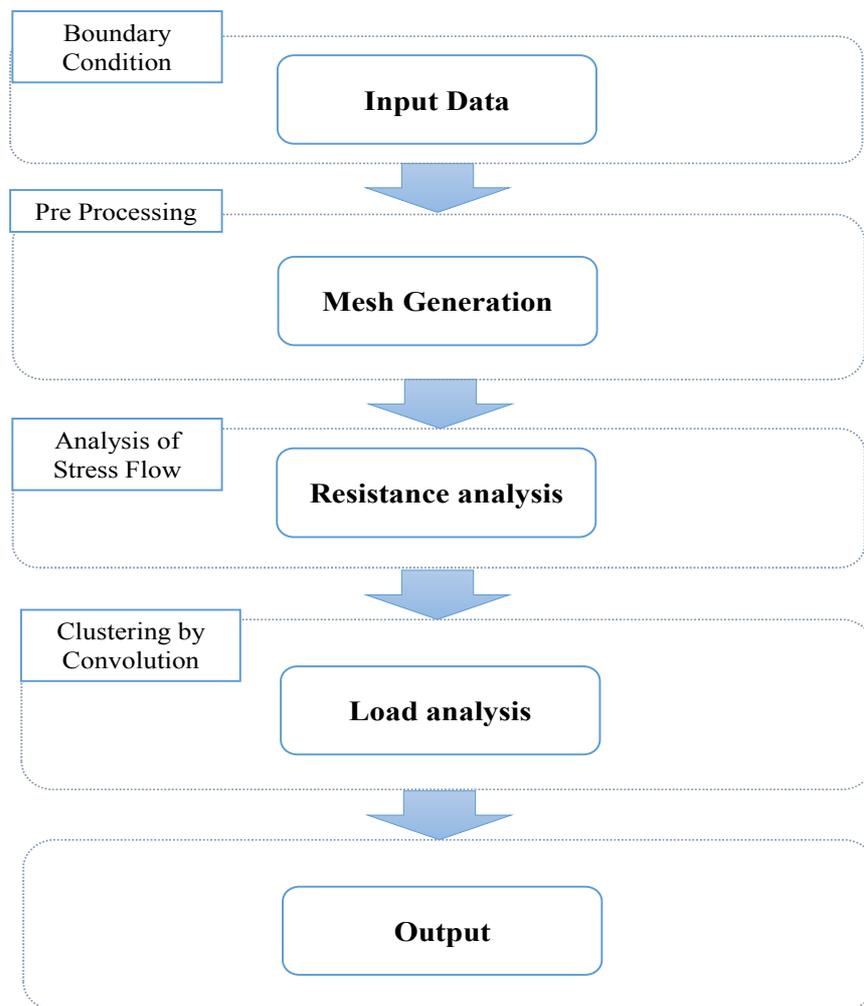


Fig. 5.7 System flow of CPC in stress path

우선 해석하고자 하는 영역에 작용하는 외력에 대해서 각 지점의 반력을 구한다. 그리고 불규칙 삼각망을 이용해서 연속체를 절점과 연결선으로 구성된 네트워크 구조 형태로 변환한다. 각 지점에서 작용하는 반력

의 흐름에 따라서 각 부재의 반력을 구하고 실제 외력의 흐름을 연산한다. 각 부재가 가지고 있는 저항에 외력의 흐름이 확률적으로 주어진다. 그러면 각 부재는 저항과 외력이 합산된 축력이 계산될 것이고 인장력이 큰 부재를 이용하면 응력 흐름을 분석할 수 있다.

Fig. 5.2와 같이 단순보에 외부 하중이 작용하면 반력에 의해서 각 부재의 저항이 결정되고 분포하중에 따라서 각 부재의 축력을 결정하면 인장력의 발현 확률이 나타난다. 모든 부재의 반력과 하중의 분포는 정규 분포이며 분산은 1로 동일하게 가정하였다.

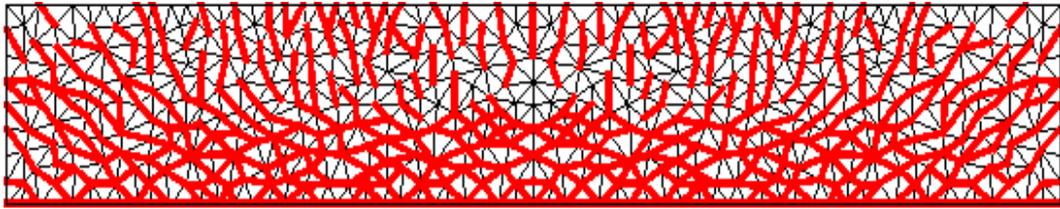


Fig. 5.8 Stress flow using CPC model with  $W_{positive} = 30\%$

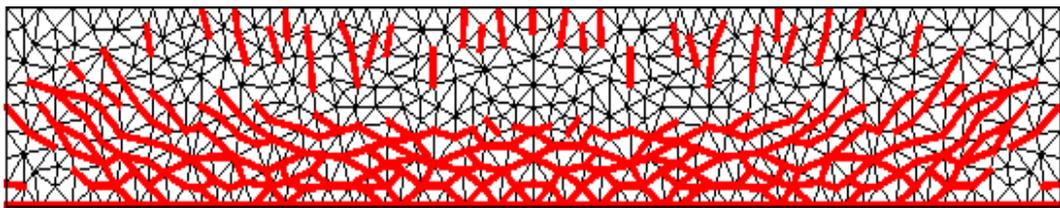


Fig. 5.9 System flow using CPC model with  $W_{positive} = 70\%$



Fig. 5.10 Theoretical result(Nawy, E. G., 1990)

인장력은 Fig. 5.8, Fig. 5.9와 같이 발견되며 균집화의 기준을 제한하였을 때 발견 정도가 구분된다. Fig. 5.9는 발견 확률을 70%로 제한했을 때이며 Fig. 5.10의 해석적인 방법을 통한 해석과 응력 흐름이 거의 일치하는 것을 확인할 수 있다. 단, 구조 중앙의 상위 부분에 인장력이 나타나는 것은 하중의 흐름과 비교해서 반력의 흐름이 적당하게 해석되지 않은 것으로 방향코사인을 이용하는 흐름 모델이 정교화되면 실제 응력 흐름과 보다 유사한 결과를 나타낼 수 있을 것으로 기대된다. 본 모델은 불균등한 요소로 구성된 연속체를 이산화 했을 때 확률적으로 해석할 수 있는 좋은 도구가 될 수 있을 것이다.

## 5.2 뼈대 구조의 파괴 모드 해석

뼈대 구조는 수평, 수직 부재와 접합부로 구성된 구조로 구조 신뢰성 해석을 통하여 외부 하중에 대한 파괴확률을 산정할 수 있다. 뼈대 구조를 네트워크로 치환하고 외부 하중을 흐름으로 가정하면 각 부재들과 접합부의 파괴가 어떻게 진행되는지 균집화 방법을 이용해서 모의할 수 있다. 응력 흐름을 균집화 하는 방법의 적용성을 검토하기 위해 Fig. 5.11과 같은 예제를 이용하였다. 본 예제는 경간과 층수가 각각 2단계인 뼈대 구조로 외부 하중이 구조물의 수평, 수직의 방향으로 각각 작용한다. 하중과 재료에 대한 확률적 분포는 Table 5.1과 같다. Palle 등은 확률적 분포를 가지는 뼈대 구조의 구조 신뢰성 해석을 위해서 잠재 힌지를 정하고 기본 파괴 모드를 기반으로 Beam, Sway, Joint 모드의 파괴확률을 각각 산정한다 (Palle and Murotsu,1986). 본 장에서는 뼈대 구조를 네트워크로 치환하고 외부 하중의 흐름에 대한 확률적 균집화 분석을 이용해서 절점들의 파괴 확률을 기반으로 Beam, Sway, Joint 모드의 파괴확률을 각각 산정하여 비교하고자 한다.

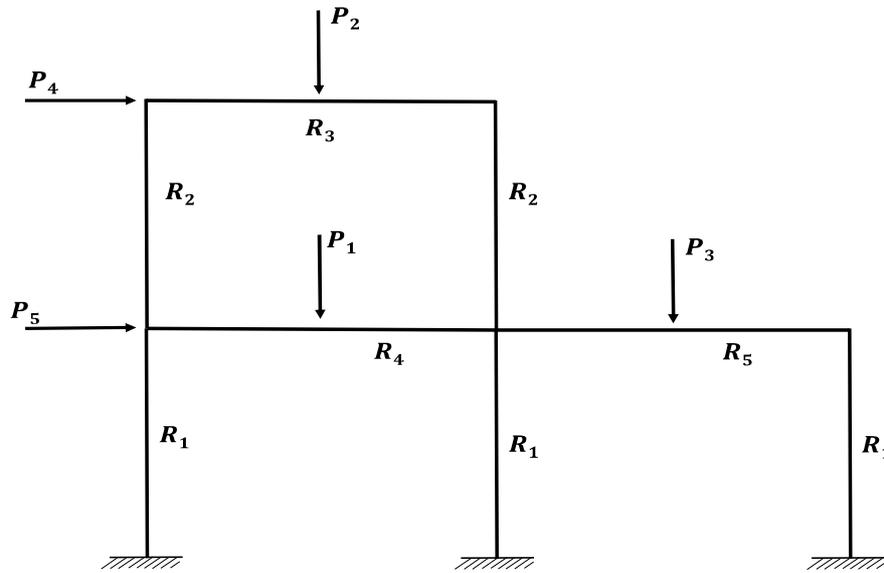


Fig. 5.11 Example of portal frame with concentrated load

Table 5.1 Boundary Condition in portal frame

Variables	Expected values	Coefficients of variation
P1	169kN	0.15
P2	89kN	0.25
P3	116kN	0.25
P4	31kN	0.25
P5	124kN	0.15
R1	95kN	0.15
R2	95kN	0.15
R3	122kN	0.15
R4	204kN	0.15
R5	163kN	0.15

### 5.2.1 뼈대 구조에서 CPC 구현

뼈대 구조 해석을 위한 균집화 방법론은 Fig. 5.12와 같이 진행된다. 먼저 뼈대 구조를 네트워크로 치환한다. 보를 절점과 연결선으로 구분하고 각 접합부를 절점으로 하면 절점과 연결선으로 구성된 네트워크가 된다.

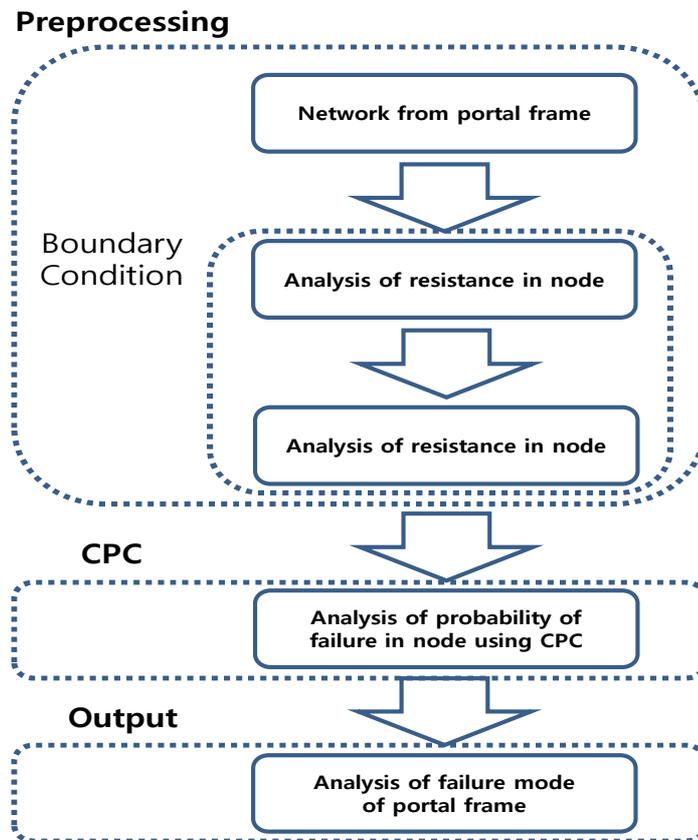


Fig. 5.12 System flow of CPC in portal frame

그리고 Table 5.1과 같이 각 부재들은 고유의 저항을 가지는데 각 절점들의 저항은 위치하고 있는 부재의 저항에 따라서 결정된다. 먼저 부재 위에 위치한 절점은 식 (14)와 같이 부재의 저항에 한 부재에 놓여있는 절점들의 개수를 나눈 값을 가진다.

$$R_i = \frac{R_{beam}}{n} \quad (14)$$

$R_i$  : resistance of node in beam  
 $R_{beam}$  : resistance of beam  
 $n$  : the number of node in beam

그리고 접합부에 존재하는 절점은 식(15)와 같이 연결된 절점들의 저항의 평균을 가진다.

$$R_j = \frac{\sum_{k=1}^n R_k}{n} \quad (15)$$

$R_j$  : resistance of node in joint  
 $n$  : the number of node linked node j

단, 외부 하중에 직접적으로 반응하는 절점은 보의 저항 전체를 가진다. Fig. 5.13과 같이 외부에서 주어진 하중  $P_a$ 는 절점에 작용하지만 보 전체에 영향을 미치며 해당 절점의 파괴 이후 작용하는  $P_b$ ,  $P_c$ 는 인접한 절점들에 단계적으로 영향을 미친다.

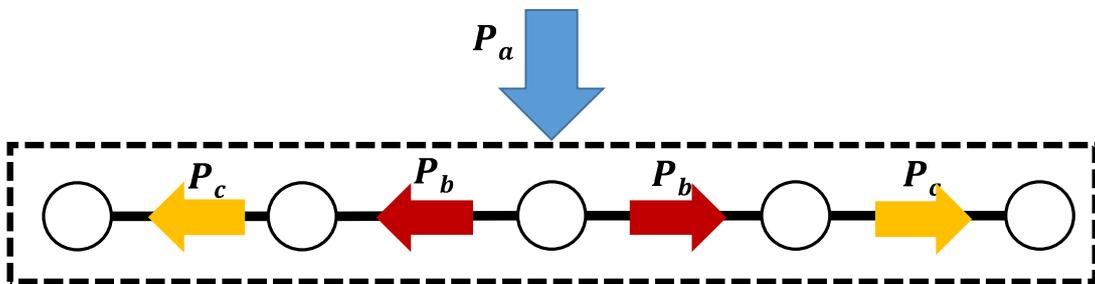


Fig. 5.13 Mechanism of stress path in portal frame

확률적 재규격화 방법론을 적용하기 위한 전처리 과정이 완성되면 외부 하중의 흐름에 따른 절점들의 반응을 분석하고 균집화 한다.

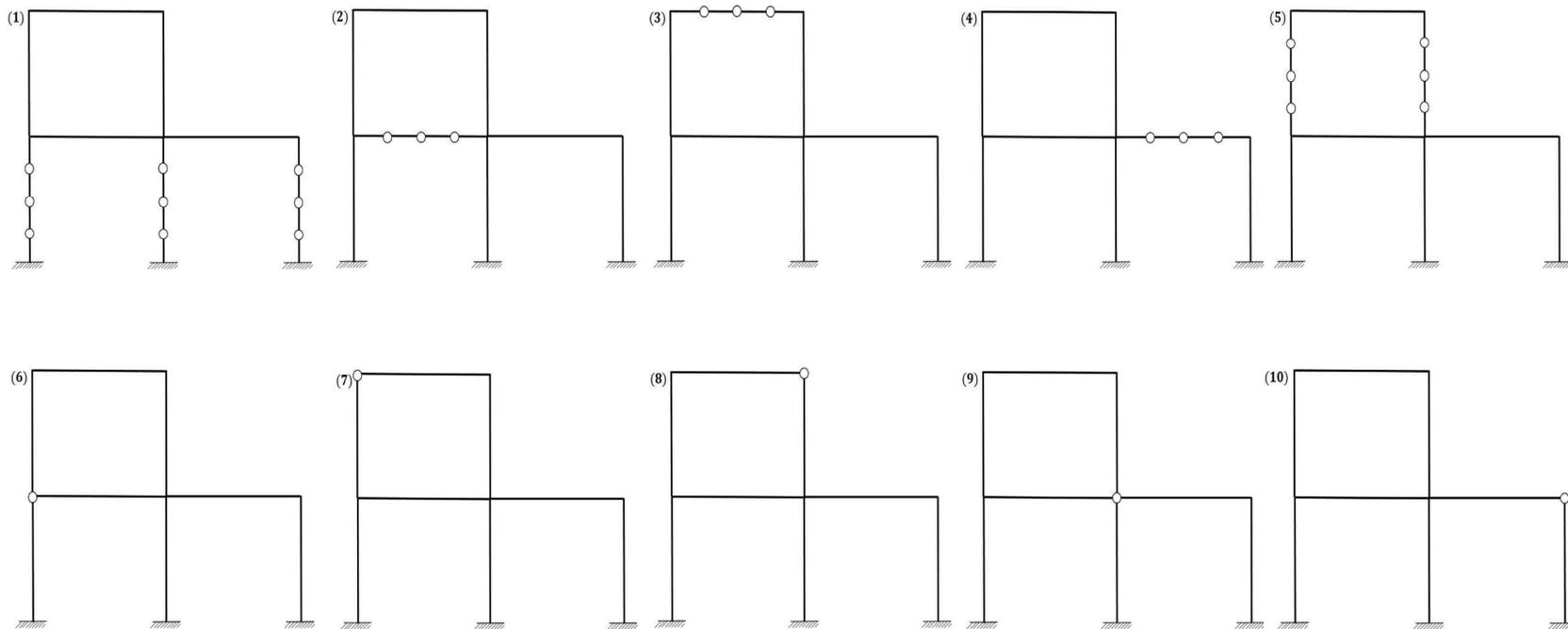


Fig. 5.14 Set of Fundamental Mechanisms with geometry, loading and potential yield hinges

외부 하중에 대한 부재력을 이용하여 하중과 저항을 단계적으로 연산하면 각 절점들의 파괴확률을 산정할 수 있다. 그리고 각 하중에 대한 반응을 합산하여 절점들의 파괴확률을 합산한다. Fig. 5.14는 Fig. 5.11의 뼈대구조를 네트워크로 치환했을 때 기본 파괴 모드를 나타낸 것이다. 여기서 (2), (3), (4)는 Beam 모드 파괴, (1), (5)는 Sway 모드 파괴, (6), (7), (8), (9), (10)은 Joint 모드 파괴로 분류하여 각 모드별 파괴확률을 산정할 수 있다.

### 5.2.2 뼈대 구조의 네트워크 치환

네트워크 균집화 모델에 적용하기 위해서 Fig. 5.15와 같이 뼈대 구조물을 절점과 연결선을 가진 네트워크로 치환한다. 한 경간에 절점은 총 5개이며 총 32개의 절점과 32개의 연결선을 가진 네트워크로 구성된다.

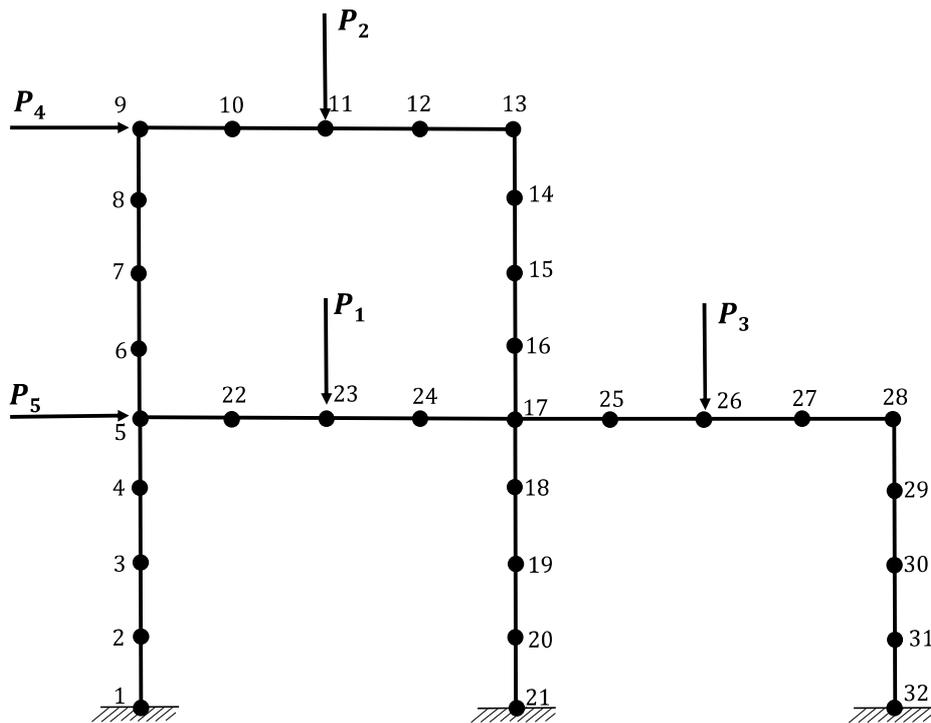


Fig. 5.15 Network model for the portal frame

네트워크 외부의 하중은 연결된 절점들을 통해서 연쇄적으로 전달되며 각 절점은 외부 하중에 대해서 파괴 확률을 가진다. 같은 경간에 있는 절점들은 거리에 따라서 경간의 저항의 평균값을 나눠가지고 분산은 동일하다. 그리고 서로 다른 경간을 연결하는 Joint 부분의 절점들은 접해있는 경간들의 평균값을 가지며 외부의 하중은 흐름과 같이 전달되며 연결되어있는 절점들로 나누어져 흘러간다.

### 5.2.3 파괴 모드 해석

32개의 절점으로 구분되어 있는 Fig. 5.15의 구조를 외부 하중에 따라서 응력 흐름을 분석하면 Fig. 5.16과 같이 나타난다. Fig. 5.16의 (1)에서 (5)까지는 외부 하중에 대한 각 절점의 응력 흐름을 나타내며 Fig. 5.16의 (6)에서 (8)까지는 지점의 반력에 대한 응력 흐름을 도식화한 것이다. 절점 간의 응력에 대한 연산 없이 네트워크에서 흐름에 대한 절점의 반응을 확인하고 균집화하는 방법론을 적용하는 것이기 때문에 반력도 외부 하중과 같이 경계조건으로 하고 응력 흐름을 모의하였다. 외부 하중에 대하여 각 지점의 반력을 구하고 반력에 대한 응력 흐름을 구할 수 있고 각 하중의 경우를 모두 합하면 각 절점들의 파괴확률이 산정된다.

응력 흐름을 분석하면 각 하중을 받은 절점들을 중심으로 연결성이 좋은 절점들이 영향을 받는다. 하중의 흐름은 절점의 저항을 거치면서 점점 작아지고 하중을 받은 절점과 연결성이 약한, 원거리에 위치한 절점들이 받는 영향은 미소하다. 예를 들어서 Fig. 5.16의 (1)의  $P_1$ 의 영향에 대한 접합부에 위치한 5번 절점의 파괴확률은 23번 절점과 22번 절점, 5번 절점의 저항과 단계적으로 컨볼루션 연산을 하여 식 (16)와 같이 구할 수 있다.

$$PDF_{P_5(x)} = conv(conv(PDF_{P_1(x)}, PDF_{R_{23}(x)}), PDF_{R_{22}(x)}), PDF_{R_5(x)}) \quad (16)$$

접합부에 위치한 절점들은 joint 모드 파괴에 영향을 미치며 경간에 위치한 절점들의 파괴는 beam, sway 모드 파괴와 관계있다. 각 절점들의 파괴 확률을 합산하여 각 파괴모드의 확률을 구할 수 있다.

뼈대 구조의 기본 파괴 기작을 크게 beam 모드, joint 모드, sway 모드로 분류할 수 있는데 1, 2, 3, 4, 6, 7, 8, 10, 11, 12, 14, 15, 16, 18, 19, 20, 29, 30, 31번 절점의 파괴는 beam 모드, 5, 9, 13, 17, 28번 절점의 파괴는 joint 모드, 1, 21, 22, 23, 24, 25, 26, 27, 32번 절점의 파괴는 sway 모드의 파괴라고 한다. 그리고 Fig. 5.14의 각 응력 흐름에 대한 절점의 반응을 합산하여 각 모드별 파괴 확률을 구하면 Table 5.2와 같다.

Table 5.2 Comparison of failure probability in CPC and the existing method

Method	MCS <sup>1</sup>	Murotsu <sup>2</sup>	CPC		
			65 nodes	32 nodes	17 nodes
Beam $P_f$	0.04666	0.04567	0.04958	0.04941	0.05001
Joint $P_f$	0.00001	-	0.00912	0.00975	0.01132
Sway $P_f$	0.02812	0.02808	0.02910	0.02924	0.02955
Total $P_f$	0.07480	0.07375	0.08780	0.08840	0.09088

1) MCS : Monte-Carlo Simulation

2) Murotsu : Palle and Murotsu, 1986

확률적 특성을 가지는 뼈대 구조의 파괴 확률을 산정하기 위해 Table 5.1의 분포와 같이 가정하고 1,000,000번의 시행으로 각 모드의 파괴 빈도를 산정하는 몬테카를로 시뮬레이션 기법(Monte-Carlo Simulation, MCS)을 적용하였다. 그리고 잠재 한지를 이용하여 기본 붕괴 모드들의 파괴 확률을 합산한 Palle와 Murotsu의 연구 방법을 적용한 결과와 한경간의 절점 수를 각각 9개, 5개, 3개로 나눠서 적용한 값과 비교하였다.

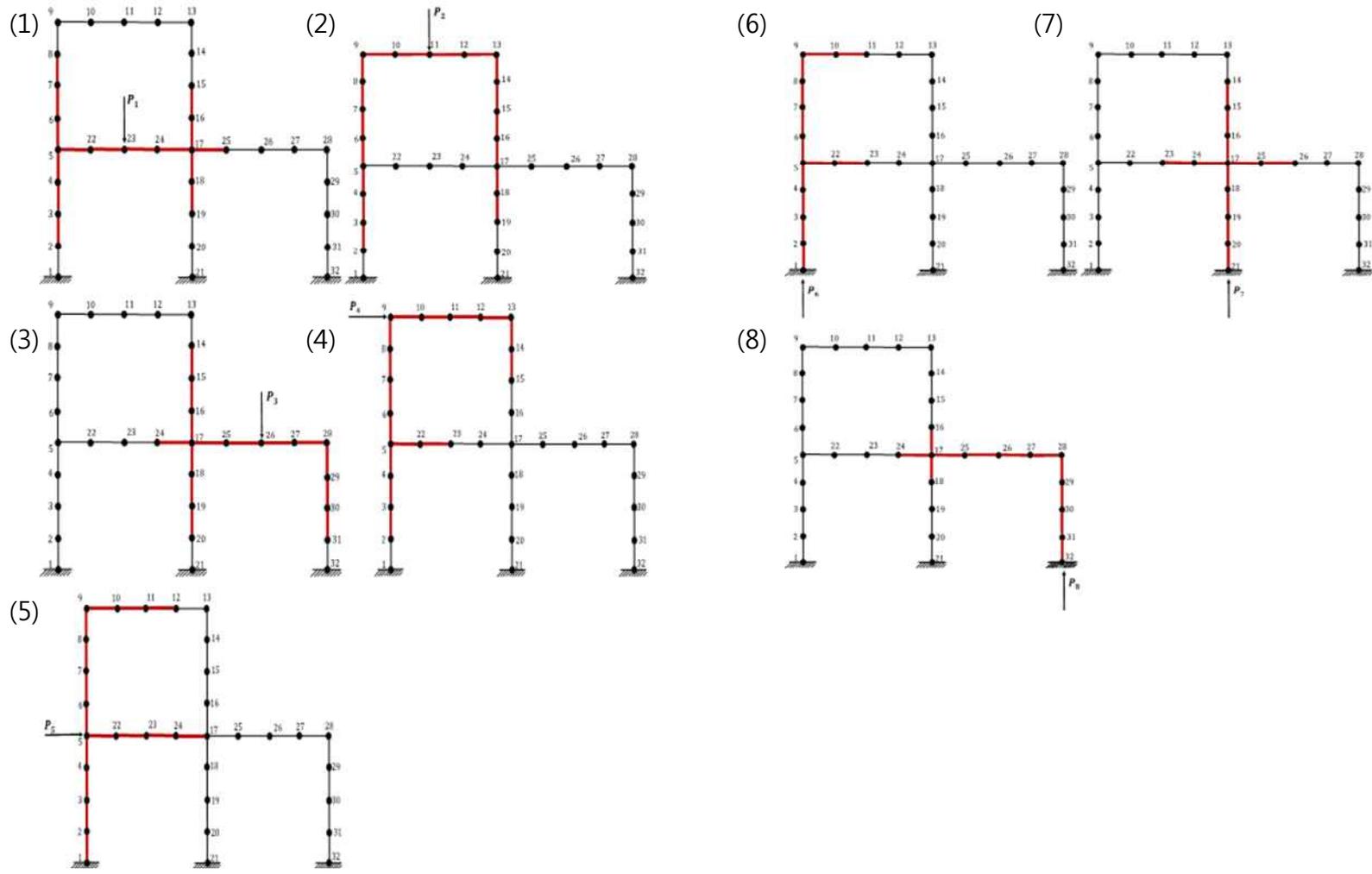


Fig. 5.16 Stress path of portal frame using CPC model

Beam 모드와 Sway 모드에서는 네트워크 균집화 방법론과 MCS, Murotsu 등의 방법이 큰 차이를 보이지 않았지만 Joint 모드에서는 큰 오차가 발생한다. 네트워크 균집화 방법론에서는  $P_4$ ,  $P_5$ 의 외부 하중이 주어지는 시작 절점이 Joint 부분이며 17번 절점의 경우 네트워크 구조에서 가장 취약한 부분이기 때문에 Beam 모드와 Sway 모드와 비교할 때 작지만 MCS와 Murotsu 등의 Joint 모드 해석 결과와 비교할 때는 큰 파괴 확률이 산정되었다. 절점 부분에 대한 분석적인 방법을 통한 보완이 필요할 것으로 판단된다.

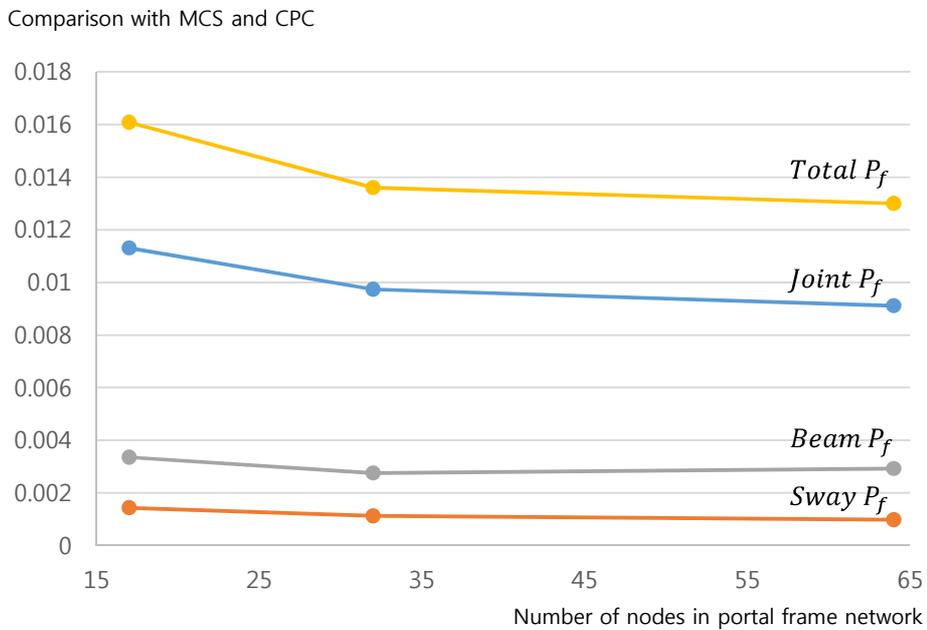


Fig. 5.17 Comparison with MCS and CPC

in the number of nodes in portal frame network.

그리고 Fig. 5.17의 결과를 보면 한 경간의 절점 수를 증가시킬수록 MCS를 이용해서 얻은 결과와 오차가 작아졌다. 그러나 절점 수가 줄어들면 계산의 효율성도 향상되기 때문에 네트워크 구조로 치환할 때 적절

한 절점 수의 범위를 산정할 필요가 있다.

네트워크 군집화를 뼈대 구조물에 적용하여 응력 흐름을 모의하는 방법은 MCS와 비교해서 Joint 모드에서 오차가 있지만 Beam, Sway 모드에서는 95% 이상의 신뢰도를 가진다. 그리고 두 방법의 계산량을 비교해보면 MCS는 정확한 결과를 얻기 위해서 1,000,000번의 시행시 파괴발생빈도수를 비교해서 값을 산정하지만 CPC의 경우 각 외부 하중에 대해서 뼈대 구조의 연결선 수만큼 컨볼루션을 수행하고 합산하여 각 절점의 파괴확률을 산정하기 때문에 보다 효율적이라는 장점을 가진다.

### 5.3 확산 네트워크의 추출과 취약성 분석

허브 네트워크는 위상적인 관계를 주로 모델링하며 무작위 네트워크는 기하학적인 관계를 모의한다. 확산 양상에 적용하면 속도가 빠른 매개체를 이용한 확산은 주로 허브 네트워크로 구축이 되며 상대적으로 속도가 느린 매개체를 이용한 확산은 무작위 네트워크로 구축된다. 확산 문제에서 속도나 매개체가 다른 네트워크가 혼재되어 있을 때 이를 추출하고 추출된 네트워크에서 확산 차단을 위하여 각각 다른 방법을 적용하여 효과적으로 대응할 수 있다.

#### 5.3.1 대상 자료

Fig. 5.18 (a)는 영국의 노스요크셔 지역의 1,419개의 축산농가 및 축산관련업체의 구제역 확산 관계를 나타낸 네트워크이다. 영국은 구제역과 HPAI와 같은 가축전염병을 관리하기 위해서 농가들의 업체 이용 관련 자료를 식품환경부에서 총괄적으로 관리한다 (Sema et al., 2011; Kao et al., 2006; Robinson et al., 2007).

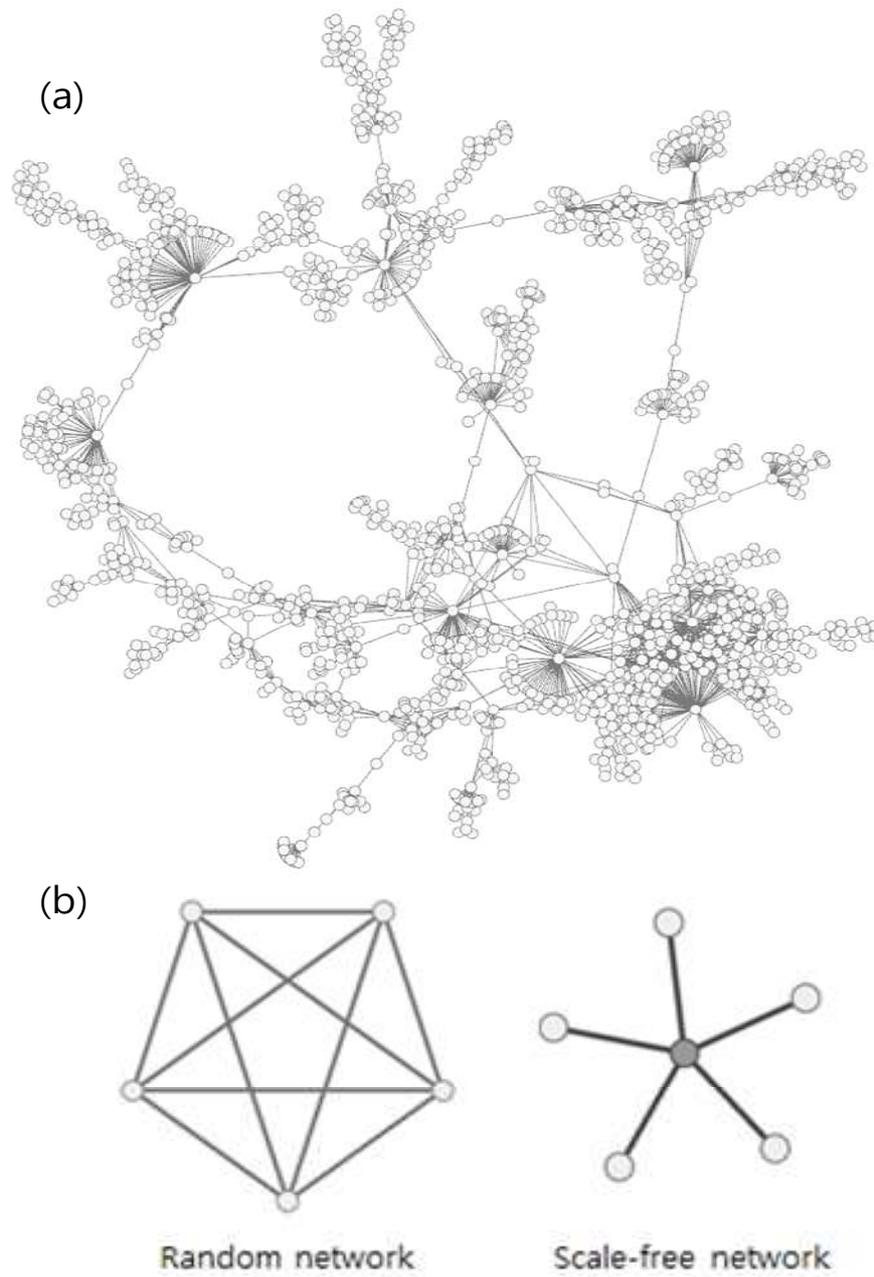


Fig. 5.18 FMD network and structure of FMD network

소지역의 농가들은 서로 인접하기 때문에 공기나 이웃농가들을 왕래하는 농장관계자를 통한 확산 가능성이 있고 서로 연결된 무작위 네트워크를 이룬다. Fig. 5.18 (b)의 무작위 네트워크와 같이 각 농가들은 해당 지역의 농가들만큼의 연결선 수를 가진다. 그리고 각 지역의 축산관련업체를 방문하는 농가들은 업체를 중심으로 허브 네트워크를 이룬다. 모든 농가들은 축산관련업체를 이용하고 있으며 일반적으로 가까운 지역의 업체를 방문한다. 다른 지역의 농가와 축산업체를 공유하는 농가는 서로 연결되고 다른 농가들보다 연결선을 많이 가진다. 그래서 Fig. 5.18 (b)의 허브 네트워크와 같이 다른 지역과의 왕래가 잦은 업체 및 농가들을 기준으로 허브 절점들이 구축된다. 같은 지역에 인접한 농가들을 연결하는 선과 도살장이나 타지역의 농가를 연결하는 선은 바이러스의 전달 매개체가 다르기 때문에 성질이 다른 두 네트워크이다. 따라서 Fig. 5.18 (a)의 구제역 확산 네트워크는 연결선 수의 분포 특성이 다른 무작위 네트워크와 허브 네트워크가 혼재되어 있는 구조이다.

### 5.3.2 확산 네트워크의 추출

Fig. 5.18 (a)의 네트워크를 대상으로 감마분포 시뮬레이션을 하면 Fig. 5.19 와 같은 분포가 나오며 형상계수는 4.17, 규모계수는 1.54이다. 각 지역을 기준으로 농가들이 군집되어 있기 때문에 4~6개 정도의 연결선을 가진 절점들이 다수를 차지하나 다른 지역과 연결되거나 업체들을 추상화한 절점들은 그 이상의 연결선을 가진다. 감마분포의 형상계수와 규모계수를 이용해서 추정한 임계치는 0.39이며 연결선의 값이 같기 때문에 연결 중심성에 따라서 절점의 위상을 나타낼 수 있고 임계치에 따라 1419개의 절점 가운데 상위 553개의 절점을 허브 네트워크의 절점으로 선정할 수 있다.

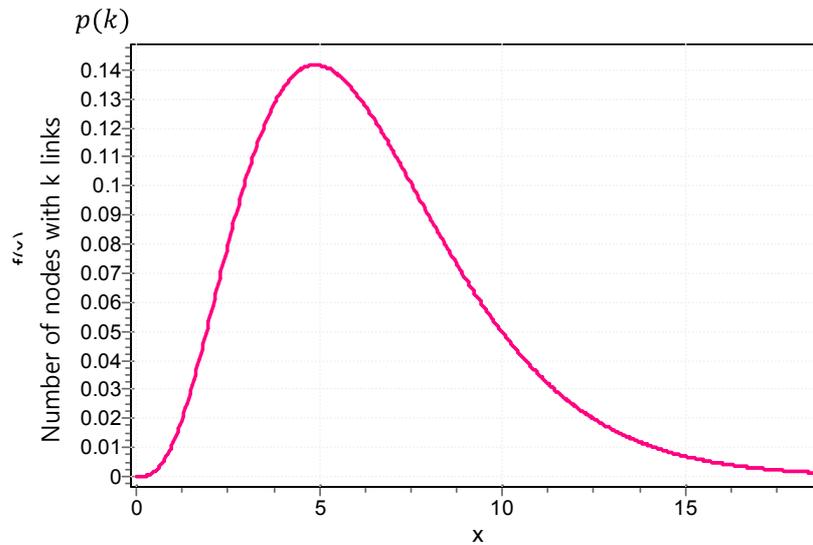


Fig. 5.19 Connectivity distribution in FMD network

### 5.3.3 확산 네트워크의 취약성 분석

임계치에 따라서 중심성이 높은 상위 553개의 절점을 가진 허브 네트워크가 추출되고 연결 중심성을 이용해서 주요 허브 절점들을 크기별로 구분하면 Fig. 5.20과 같다. 구제역 네트워크 전체 구조에서 바이러스의 접촉 빈도가 높기 때문에 허브 절점들은 가장 취약한 지점이라고 볼 수 있다. 허브 절점이 감염되면 많은 절점에게 영향력을 가지기 때문에 연쇄적으로 감염이 전달될 수 있고 네트워크 구조가 붕괴될 수 있다.

따라서 중추적인 역할을 하는 허브 절점들을 우선적으로 차단하면 소지역뿐만 아니라 타지역으로 확산을 차단할 수 있는 효율적인 방역대책을 수립할 수 있다. 사전에 중심성이 높은 절점들을 중심으로 예방할 수 있는 대책을 지원한다면 구제역이 발병하더라도 광역적으로 확산이 되는 것을 차단할 수 있다.

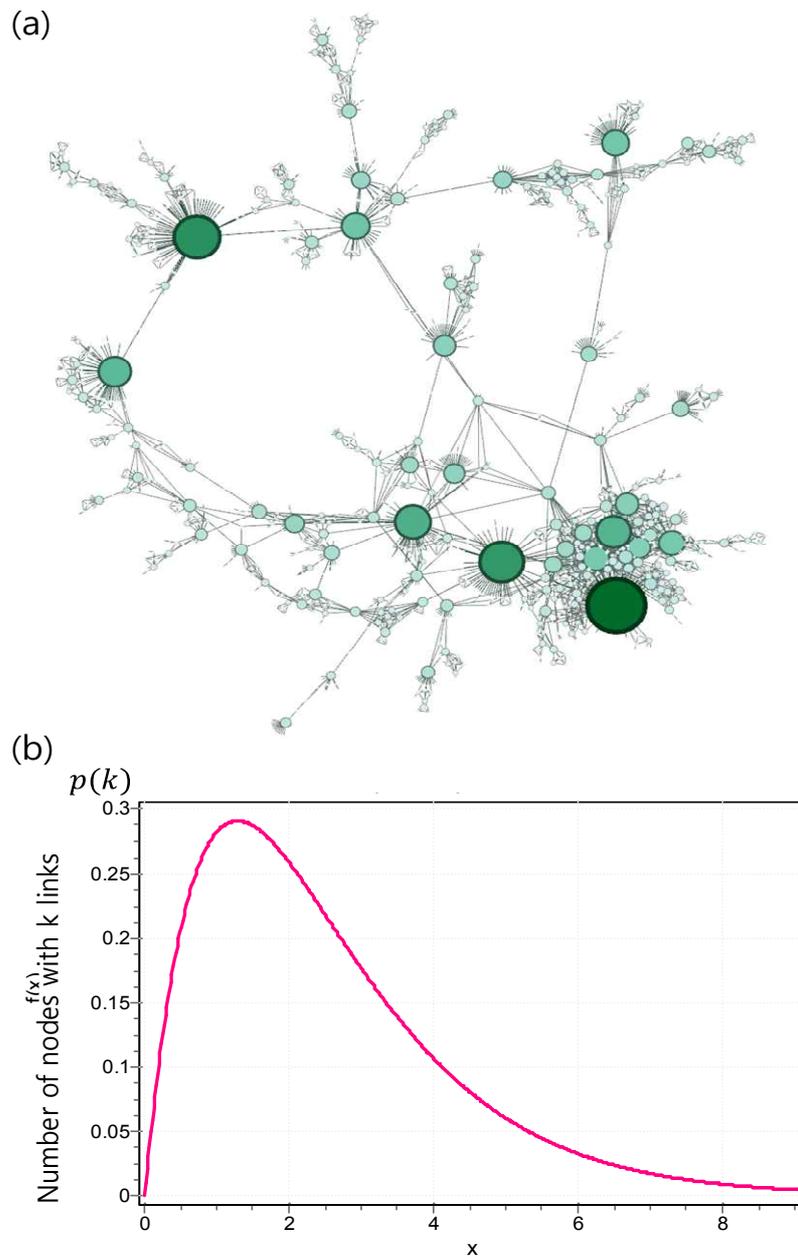


Fig. 5.20 Extraction of hub network from FMD network and Connectivity distribution in hub network

허브 네트워크를 추출하고 남은 절점으로 구축된 네트워크는 866개의 절점으로 구성된 무작위 네트워크이다. 각 지역의 농가들과 연결되어 있는 무작위 네트워크의 모임이기 때문에 허브 네트워크의 허브 절점을 이용한 확산 차단 방법과는 다르게 Fig. 5.21과 같이 공간 차단(Ring Vaccination)으로 확산에 대처할 수 있다 (Muller et al., 2000; Keeling et al., 2003).

주요한 절점들을 이용해서 기하학적인 요소와 관계없이 확산되지 않고 기하학적으로 근접한 절점들을 대상으로 확산이 진행되기 때문에 발병 지역 일정 거리를 통제하는 방법으로 권역을 설정하고 방제하면 점진적인 공간을 통한 확산에 효과적으로 대처할 수 있다.

확산을 모의하는 네트워크에서 무작위 네트워크와 허브 네트워크는 확산 진행 양상이 다르다. 거리에 관계없이 허브 절점을 중심으로 불균등하게 구축된 허브 네트워크는 장거리로 확산하는 매개체를 해석하는데 유용하며 무작위 네트워크는 각 소지역을 중심으로 거리에 따라서 균등하게 확산되는 매개체를 해석하기에 적합하다. 따라서 매개체가 혼재되어 있는 확산 네트워크에서 허브 네트워크와 무작위 네트워크를 각각 추출하고 각 특성에 맞게 네트워크를 해석하면 확산에 대한 최적의 방역대책 수립이 가능할 것으로 기대된다.

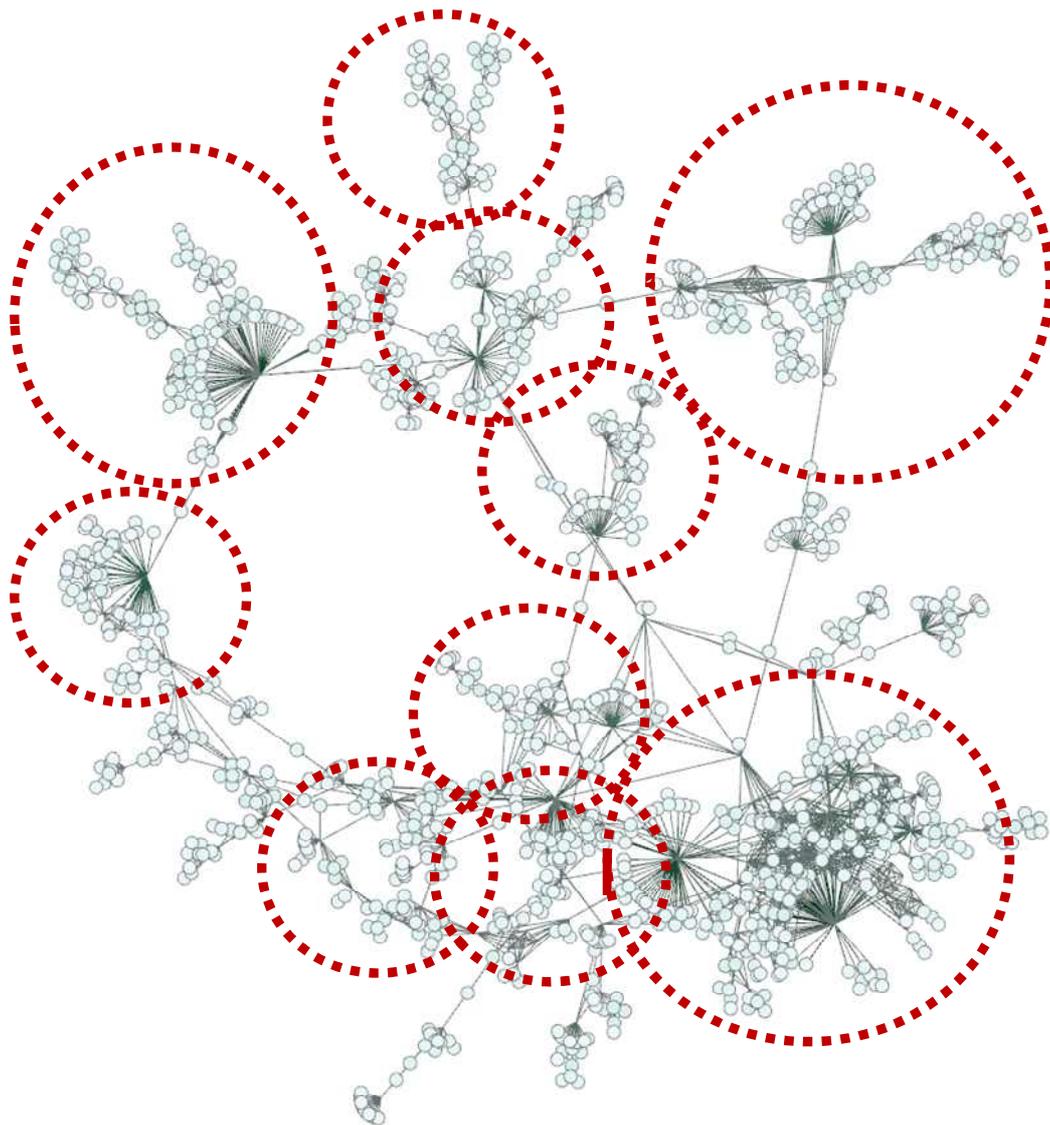


Fig. 5.21 Ring Vaccination in random network from FMD network

## 참고문헌

1. Hinton, E., and Sienz, J., 1995, Fully stressed topological design of structures using an evolutionary procedure, *Eng. Comput.*, 12(2), 229-244.
2. Johannes Muller, Birgitt Schonfisch, Markus Kirkilionis, 2000, Ring Vaccination, *J. Math. Biol.* 41, 143-171
3. Kao, R.R., L. Danon, D.M. Green and I.Z. Kiss, 2006. Demographic structure and pathogen dynamics on the network of livestock movements in Great Britain. *Proc.*
4. Keeling, M. J., et al. 2002, Modelling vaccination strategies against foot-and-mouth disease. *Nature* 421.6919, 136-142.
5. Lee, H. J., 2008. Development of the Stress Path Search Model Using Refined Evolutionary Structural Optimization. MS. diss., Seoul National University.
6. Lee, S. Y., 2009. Equivalent Truss Model by Discretizing continuum Structure. MS. diss., Seoul National University.
7. Leu, L.-J., C. W. Huang, C. S. Chen and Y. P. Liao, 2006. Strut-and-tie design methodology for three-dimensional reinforced concrete structures. *Journal of structural engineering* 132(6): 929-938.
8. Nawy, E. G., 1990, Reinforce Concrete a Fundamental Approach Terjemahan.
9. Palle, T.C., and Murotsu, Y. 1986, Reliability Analysis of Structural Systems by the  $\beta$ -Unzipping Method, Application of Structural

- Systems Reliability Theory. Springer Berlin Heidelberg, 143–214.
10. Robinson, S.E, M.G. Everett and R.M. Christley, 2007. Recent network evolution increases the potential for large epidemics in the British cattle population. *J. R. Soc. Interface* 4(15): 669–674.
  11. Sema, N., L. Matthews, P.R. Bessell, S. WJ Reid and R.R. Kao, 2011. Generating social network data using partially described networks: an example informing avian influenza control in the British poultry industry. *BMC Veterinary Research*, 7(66).
  12. Truesdell, C. and R. Toupin, 1960. *The classical field theories*. Springer Berlin Heidelberg.
  13. Truesdell, C. and W. Noll, 2004. *The non-linear field theories of mechanics*. Springer Berlin Heidelberg.
  14. Yerry, M. A. and M. S. Shephard, 1984. Automatic Mesh Generation by the modified-octree Technique. *International Journal for Numerical Mesh in Engineering* 20(11): 1965–1990.
  15. Yoon, S. S., 2001. *An Integrated Design System for Irrigation Facilities Using Object-Oriented and Evolutionary Structural Optimization Techniques*. Ph.D. diss., Seoul National University.



## 제 6 장 결 론

### 6.1 요약 및 결론

복잡한 연결 관계를 가지는 대상을 간단하게 절점과 연결선으로 모델링할 수 있는 네트워크는 전체 구조의 특성을 대표하는 허브 네트워크와 배경과 같은 무작위 네트워크로 구분할 수 있다. 이 두 가지의 서로 다른 네트워크가 혼재되어 있는 실세계를 해석할 때 전체 구조를 대표하는 허브 네트워크를 추출할 수 있다면 해석의 효율성과 결과의 정확도를 향상시킬 수 있다.

추출된 허브 네트워크의 해석 방법은 전체 구조의 거동을 관찰하는 거시적인 분석과 구조 내 요소들의 속성을 관찰하는 미시적인 분석으로 구분할 수 있다. 거시적인 분석은 불균등한 허브 네트워크에서 절점들의 연결성을 분석하고 네트워크의 생성, 소멸과 같은 외적 변화를 관찰하는 방법이며 미시적인 분석은 경계조건이 고정된 네트워크에서 절점들의 속성 변화를 분석하는 방법이다.

본 연구에서는 혼합된 네트워크에서 허브 네트워크를 추출하고 추출된 허브 네트워크를 포괄적으로 분석하는 방법론을 개발하였다. 그리고 이를 확산 및 구조 신뢰성 문제에 적용하고 고찰하였다. 2장부터 5장까지 내용을 요약하면 다음과 같다.

- 혼합 네트워크는 무작위 네트워크와 허브 네트워크가 혼재되어 있다. 네트워크는 절점과 연결선 수의 관계에 따라서 연결선 수 분포라는 특성을 가지는데 이 분포를 감마분포로 나타내면 형상계수와 규모계수라는 고유의 값을 구할 수 있다. 2 장에서는 이 고유값을 이용하여 혼합 네

트위크에서 허브 네트워크를 추출할 수 있는 임계치를 산정하는 방법론을 개발하였다.

- 간선과 지선이 혼재되어 있는 교통망 네트워크에 적용하여 검증한 결과 본 연구의 방법론으로 추출한 허브 네트워크와 간선 네트워크의 연결선 수가 3.81%의 오차를 가지는 것을 확인하였다.
- 2 장에서 개발한 혼합 네트워크에서 대표성 있는 허브 네트워크를 추출하는 방법은 허브 네트워크를 해석하기 위한 전단계로 유용하게 활용될 수 있을 것이며 감마분포로 형상계수와 규모계수를 추정하는 단계를 개선하면 보다 정확한 추출 기준을 산정할 수 있는 방법으로 개선될 수 있을 것으로 기대된다.

• 추출된 허브 네트워크의 절점들은 서로 다른 중요도를 가진다. 허브 네트워크의 연결성을 분석하면 중추적인 역할을 하는 허브 절점을 탐색할 수 있다. 이 허브 절점은 수많은 절점에 영향을 미치기 때문에 외부에서 공격받을 경우 연쇄적으로 연결된 다른 절점들에게 영향을 미치고 전체 구조를 붕괴시킬 수 있다. 3 장에서는 가축전염병 확산 네트워크를 대상으로 취약성 분석을 수행하고 허브의 역할을 고찰하였다.

- HPAI 확산을 모의하는 네트워크에서 허브 농가의 차단 유무에 따라 확산 양상을 비교해보면 허브 농가를 차단했을 때 전체 28% 이상의 농가들이 감염되지 않고 보다 안전한 것을 확인하였다.
- 확산 네트워크에서 허브 절점을 우선적으로 통제하는 방법은 효율적인 확산 차단 대책으로 활용될 수 있다.

• 허브 네트워크의 절점들은 연결성에 따라서 작은 집단들로 군집화될 수 있다. 4 장에서는 절점과 연결선의 값이 확률적으로 변동하는 경우 확률론적 컨볼루션을 이용해서 각 절점들의 관계를 연산하면서 유사

한 특성을 가지는 절점들을 군집화 하는 방법론을 개발하였다.

- 임의의 네트워크를 대상으로 확률론적 군집화 방법론과 몬테카를로 시뮬레이션을 통한 방법을 각각 적용하고 비교했을 때 0.5216%의 오차를 가지며 기존의 다른 군집화 방법과 비교했을 때 군집화의 기준에 따라서 동일한 결과를 얻음을 확인하였다.
- 이 방법은 몬테카를로 시뮬레이션과 같이 1,000,000번의 시행 없이 한 번의 시행으로 각 절점의 군집화를 수행할 수 있는 방법으로 효율적이며 경계조건이 고정된 네트워크에서 내부의 속성을 분석하는 방법으로 구조물의 응력 흐름 해석 및 뼈대 구조의 구조 신뢰성 해석 등에 활용할 수 있다.
- 5 장에서는 앞서 소개된 방법론을 구조 신뢰성 문제와 확산 문제에 각각 적용하였다. 경계조건이 고정된 외력을 받는 연속체 구조물을 네트워크로 치환하고 허브 네트워크 추출 방법론과 확률론적 군집화 방법론을 적용하여 인장력의 흐름을 분석하였고 뼈대 구조에 적용하여 구조 신뢰성 해석을 통해서 파괴확률을 산정하였다. 그리고 구제역 확산 네트워크를 대상으로 허브 네트워크와 무작위 네트워크로 추출하여 각 네트워크에서 확산을 차단하기 위한 방안을 제시하였다.
- 연속체 해석에서 기존의 구조 최적화 기법의 수행지수와 비교했을 때 1.1%의 오차가 있지만 임계치를 단계별로 산정할 필요 없이 연결선 수 분포를 이용해서 효율적으로 최적 기준을 산정할 수 있었다.
- 뼈대 구조의 구조 신뢰성 해석시 네트워크 구조로 치환하여 군집화 방법론을 적용하면 몬테카를로 시뮬레이션을 통한 방법과 비교했을 때 Beam 모드와 Sway 모드에서는 95%이상의 신뢰도를 가지며 경

간의 절점수를 증가시키고 영역을 세분화할수록 더 정확한 결과를 얻음을 확인하였다.

- 이 방법은 각 외부 하중에 대한 한 번의 연산으로 파괴확률을 해석하는 방법으로 몬테카를로 시뮬레이션을 통한 해석과 비교할 때 효율적이다.
- 구제역 확산 문제에서 방역을 위해 기존의 공간 확산 차단 방법을 이용하면 전체의 약 87%의 농가를 통제해야하는 반면에 네트워크로 치환하고 허브 네트워크를 추출하면 12% 농가만의 통제로 기존의 방역 효과를 얻을 수 있음을 보였다.

본 연구는 실제 문제를 해석하는데 결과의 정확성 및 방법의 효율성을 향상시키기 위해 혼합 네트워크에서 허브 네트워크를 추출하는 방법론을 정립하였다. 그리고 추출된 허브 네트워크를 거시적인 관점의 취약성 분석과 미시적인 관점의 군집화 분석하는 방법론을 개발하였고 이를 각각 확산 문제와 구조 신뢰성 문제에 적용하였다. 네트워크를 이용하여 모델링하고 해석하는 방법은 지역시스템공학의 발생할 수 있는 문제를 해결하는데 기여할 수 있을 것으로 기대된다.

## 6.2 향후 연구

각 장의 방법론에 적용된 대상 및 고찰을 바탕으로 한계와 향후 연구에 대해서 검토하면 다음과 같다.

- 감마분포를 이용해서 허브 네트워크의 규모계수와 형상계수를 추정할 때 최빈값을 1로 가정했기 때문에 전체 구조의 특성을 대표하는 허브

네트워크를 추출하는데 한계가 있을 수 있다. 따라서 규모계수를 혼합 네트워크의 규모계수와 동일하게 가정하는 단계와 최빈값을 1로 가정하는 단계에서 보다 분석적으로 접근한다면 추출의 정확도를 향상시킬 수 있을 것이다.

- 확산 네트워크를 대상으로 취약성 분석을 수행하는데 적용한 자료의 질이 향상되면 농가 간을 연결하는 연결선에 가중치와 방향성을 줄 수 있을 것이고 보다 정확한 확산 양상을 모의할 수 있을 것으로 판단된다.
- 일방향성의 네트워크를 대상으로 적용한 군집화 방법론에서 양방향성이 가능한 확률적 연산을 개발한다면 응력 흐름에 대한 구조 해석을 수행할 때 보다 정확한 결과를 도출할 수 있을 것으로 기대된다.

각 장의 방법론은 네트워크의 시간적 변화와 내용적 변화를 포괄적으로 분석한 것으로 적용 대상의 범위가 확대되며 오차를 줄일 수 있는 방향으로 향후 연구가 발전된다면 보다 범용적인 분석 방법으로 확장될 수 있을 것으로 기대된다.



# Spreading and Structural Reliability Analysis by Extraction of Hub Network with Gamma Distribution

Hyungjin Lee

Major in Rural Systems Engineering

Department of Landscape Architecture and

Rural Systems Engineering

The Graduate School

Seoul National University

## Abstract

The system in the real world is the combination of two kinds: hub system which is composed of hub and represents the entire structure and random system which is identical in all places, such as background. If hub system representing the whole structure can be

extracted from an object, it would improve accuracy and efficiency of interpretation at the same time.

When the continuum was analysed into a simple discretized model, the network can also be classified into two: hub network with unevenly distributed nodes and random network with evenly distributed nodes. These two types of network have different characteristics and should be analysed differently as well. Therefore, this study extracted representative hub network from real problem and analysed vulnerability in a macroscopic viewpoint and detected the community in a microscopic viewpoint to problems such as diffusion and structural reliability.

Every network has unique link distribution of which the number of link of nodes is the variable. Using gamma distribution, shape factor and scale factor, the eigen values of distribution, can be calculated. Two eigen values can be used to calculate the threshold at which hub network can be extracted from the object.

Each node has different significance in accordance with link since the extracted hub network has uneven distribution. Hub node with many links has decisive influence in the volume of the entire network and thus useful to see external changes such as formation and extinction of network. This study considered the application of hub investigation to the spread and the extinction of livestock epidemics to achieve effective prevention of spread.

If the features of links in externally fixed hub network fluctuate stochastically, internal change of the entire network can be taken into account through detecting the community with convolution. This study examined efficient application of structural reliability through detecting the community of hub network which affect the destruction

in portal frame.

By applying extraction of hub network and analysing including vulnerability analysis and detecting the community respectively to structural problem and spread problem, it was possible to achieve accuracy over 98% with just link distribution analysis without calculating strain energy and volume in each stage, compared to performance index for optimal structural extraction in continuum. In addition, in the analysis the structural reliability in the portal frame, single detecting the community of each external force, not like in Monte Carlo interpretation which needed 1,000,000 arithmetic operations, resulted in beam, sway mode destruction probability with accuracy over 95%. In livestock epidemic spread problem, it was also witnessed that extraction of hub network and centrality analysis led to blockage of merely 12% of farms, so that its effect could be same as that of the existing method which blocked 87% of farms.

**Keywords : Hub network, Gamma distribution, Vulnerability,  
Clustering, Convolution, Portal frame**

*Student Number : 2008-30922*