

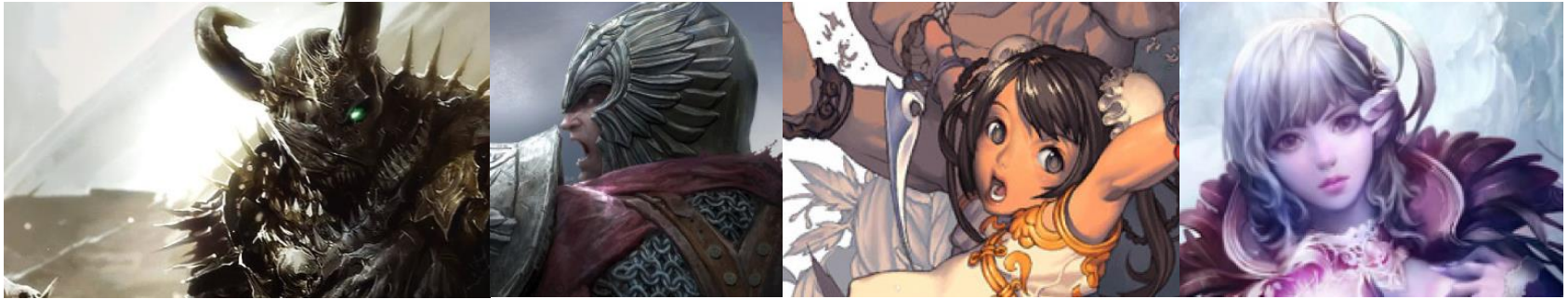
## 네트워크 분석 기법을 활용한 게임 데이터 분석

이은조

엔씨소프트

# R User Conference 2018

## 네트워크 분석 기법을 활용한 게임 데이터 분석



엔씨소프트 이은조



# Introduction

## 와츠 & 스트로가츠 논문 발표 20주년

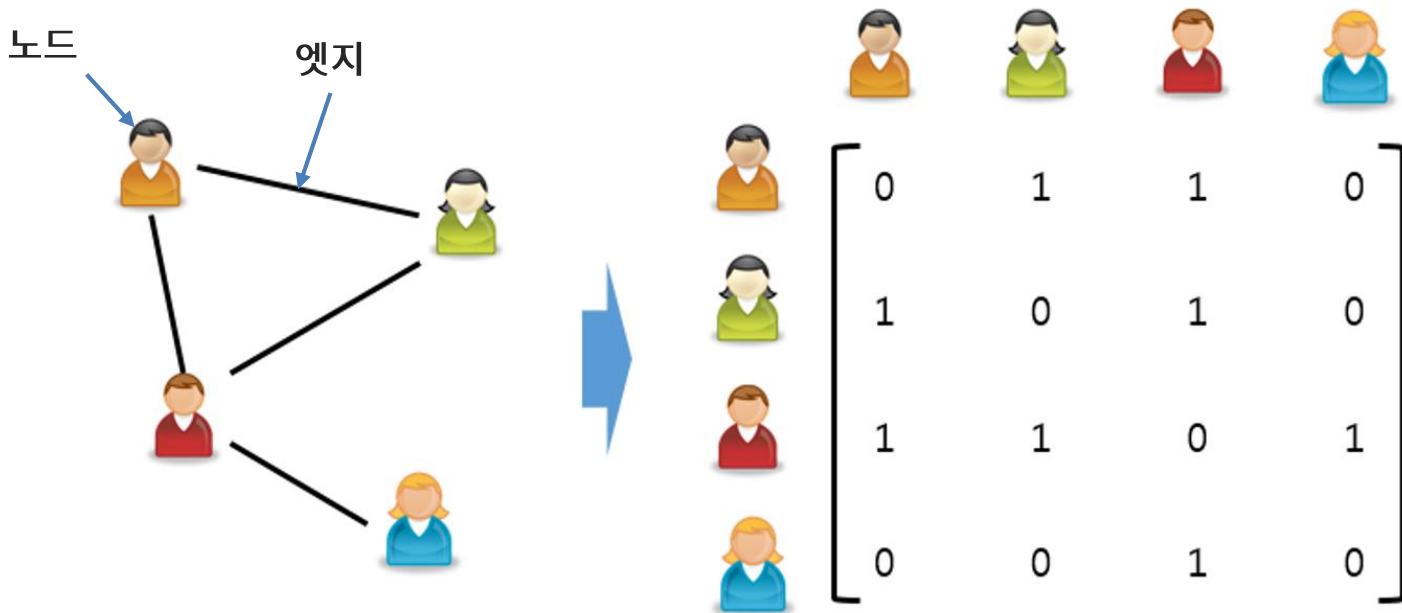
- 네트워크 과학의 시작을 알린 기념비적 논문
- 다양한 분야의 현상들이 갖는 공통점
  - ✓ 배우들의 영화 출연 관계
  - ✓ 전력선 연결
  - ✓ 예쁜 꼬마 선충의 뉴런 연결 구조

The screenshot shows the top portion of a Nature journal article page. At the top, there is a dark red navigation bar with the 'nature' logo and a 'MENU' dropdown. Below this, the article title 'Collective dynamics of 'small-world' networks' is prominently displayed. The authors, Duncan J. Watts and Steven H. Strogatz, are listed below the title. The article is identified as a 'Letter' published on '04 June 1998'. On the right side, there are statistics for '17263 Citations' and '157 Altmetric' scores. Below the title, there are sections for 'Associated Content', 'Abstract', and a sidebar with 'Sections', 'Figures', and 'References' tabs. The 'Abstract' section is currently selected and contains the following text: 'Networks of coupled dynamical systems have been used to model biological oscillators<sup>1,2,3,4</sup>, Josephson junction arrays<sup>5, 6</sup>, excitable media<sup>7</sup>, neural networks<sup>8,9,10</sup>, spatial games<sup>11</sup>, genetic control networks<sup>12</sup> and many other self-organizing systems. Ordinarily, the connection topology is assumed to be either completely regular or completely random. But many biological, technological and social networks lie somewhere between these two extremes. Here we explore simple models of networks that can be tuned through this middle ground: regular networks 'rewired' to introduce increasing amounts of disorder. We find that these systems can be highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs. We call them 'small-world' networks, by analogy with the small-world phenomenon<sup>13, 14</sup> (popularly known as six degrees of separation<sup>15</sup>). The neural network of the worm *Caenorhabditis elegans*, the power grid of the western United States, and the collaboration graph of film actors are shown to be small-world networks. Models of dynamical systems with small-world coupling display enhanced signal-propagation speed, computational power, and synchronizability. In particular, infectious diseases spread more easily in small-world networks than in regular lattices.'

# Introduction

## 네트워크 분석이란?

- 네트워크 (그래프): 노드와 엣지로 구성되는 자료 구조
- 네트워크 분석: 네트워크 구조로 이뤄진 자료에 대한 분석

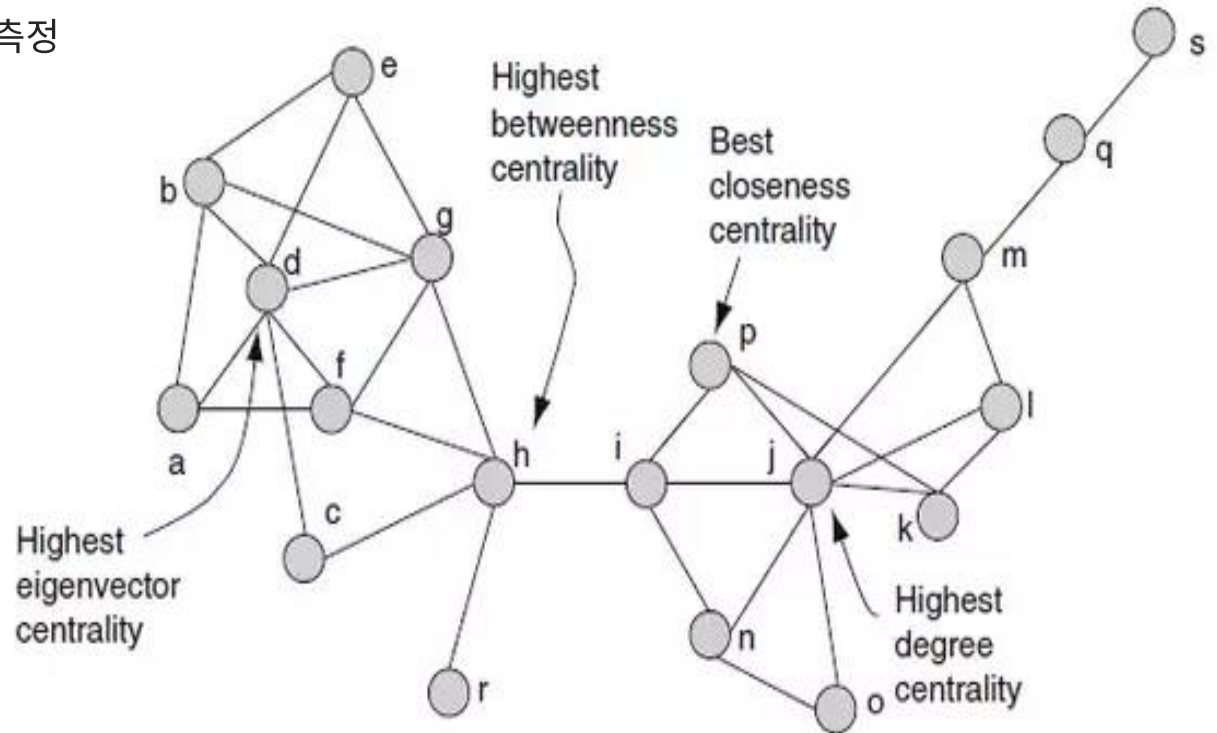


# Introduction

## 주요 네트워크 분석 기법 - Node centrality

➤ 전체 집단에서 **노드의 중요도** 측정

- ✓ Degree
- ✓ Betweenness
- ✓ Closeness
- ✓ Eigenvector



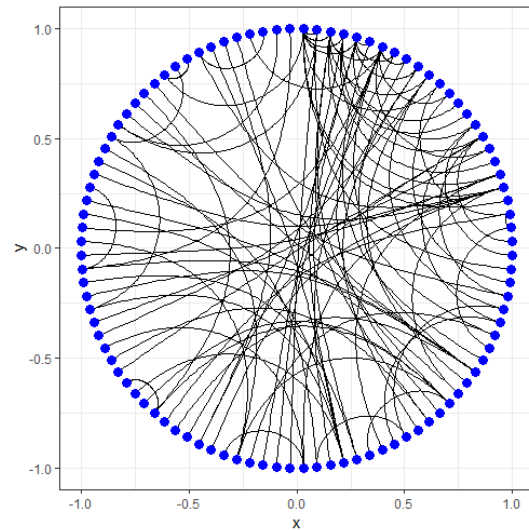
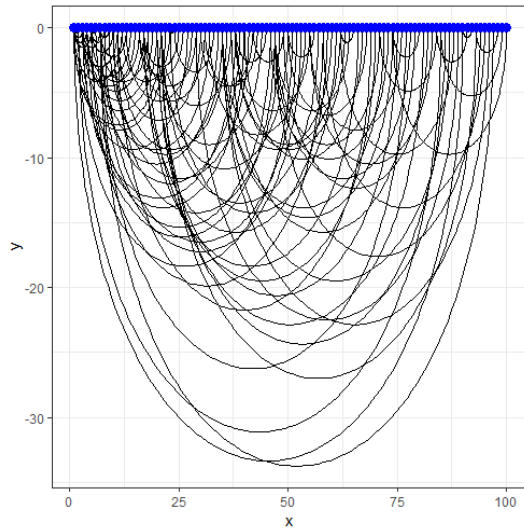
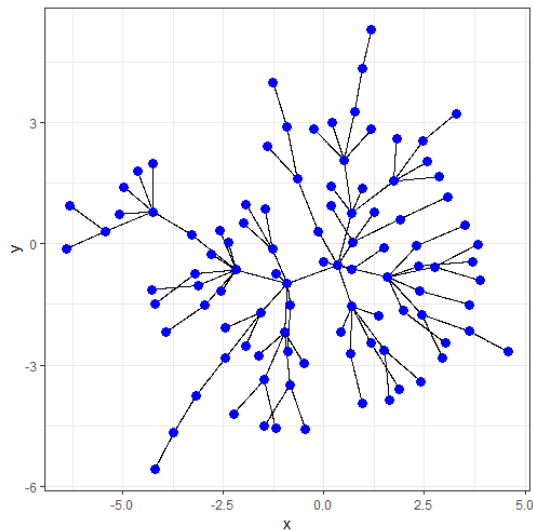
# Introduction

## 주요 네트워크 분석 기법 - Network structure

➤ 단순히 외형이 아닌 **위상 구조 파악**이 필요함

- ✓ Size
- ✓ Radius
- ✓ Clustering coefficient
- ✓ Degree assortativity
- ✓ Mean / STD of node centrality

Size	100
Radius	6
Clustering coefficient	0
Degree assortativity	-0.143
Mean / STD of degree	1.98 / 1.68
Mean / STD of betweenness	3.45 / 7.31



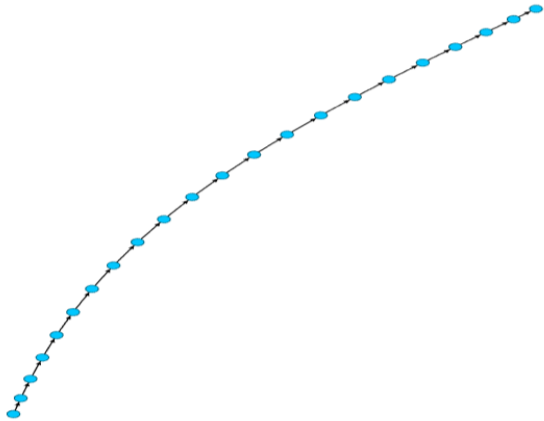
# Introduction

## 주요 네트워크 분석 기법 - Network structure

### ➤ Radius

- ✓ 접근성 (Closeness centrality) 이 가장 좋은 노드에서 가장 먼 노드까지의 거리

Large radius network



Small radius network

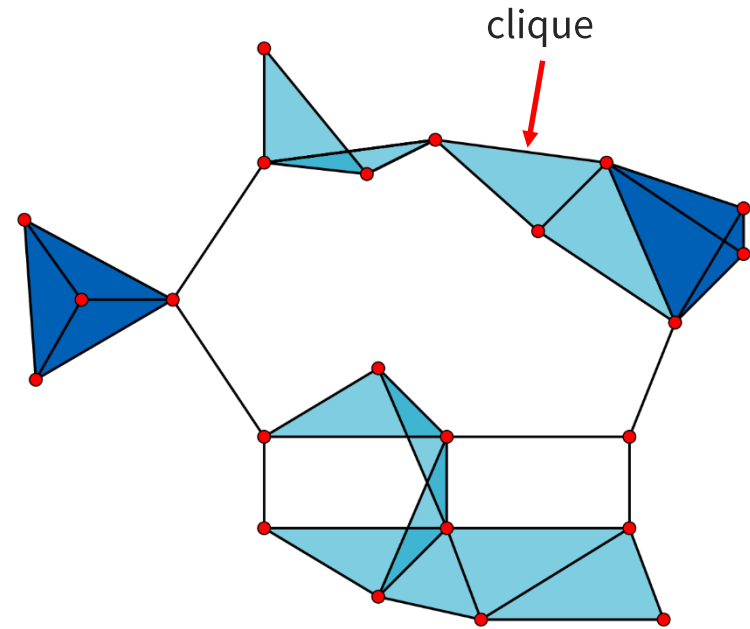
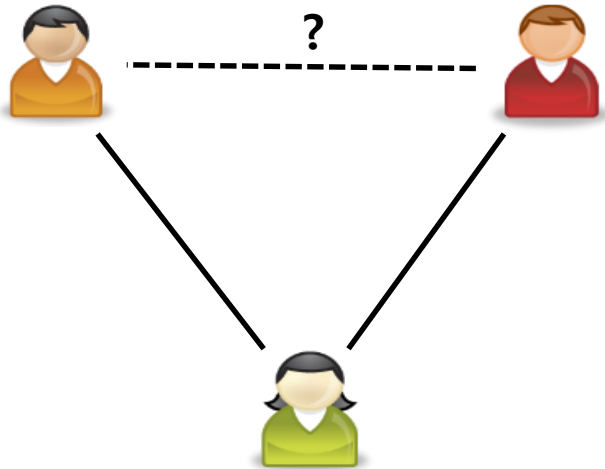


# Introduction

## 주요 네트워크 분석 기법 - Network structure

### ➤ Clustering coefficient

- ✓ 나의 서로 다른 두 친구가 서로 친구일 확률



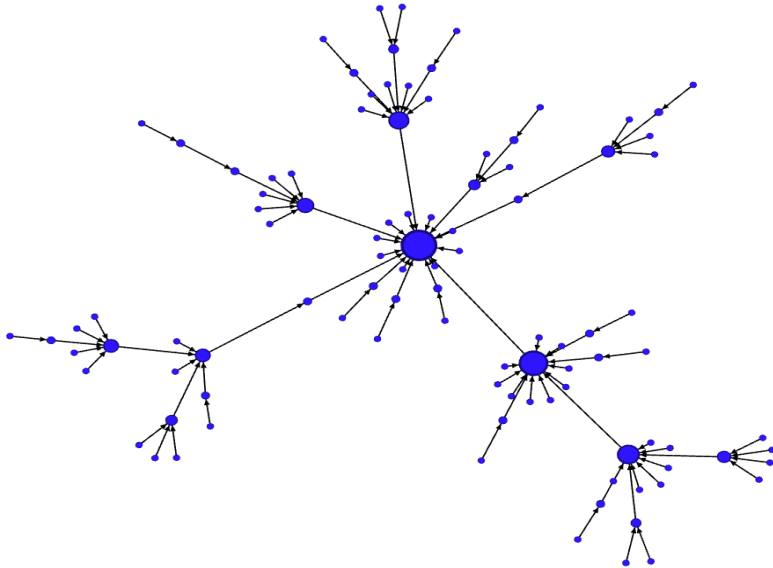


# Introduction

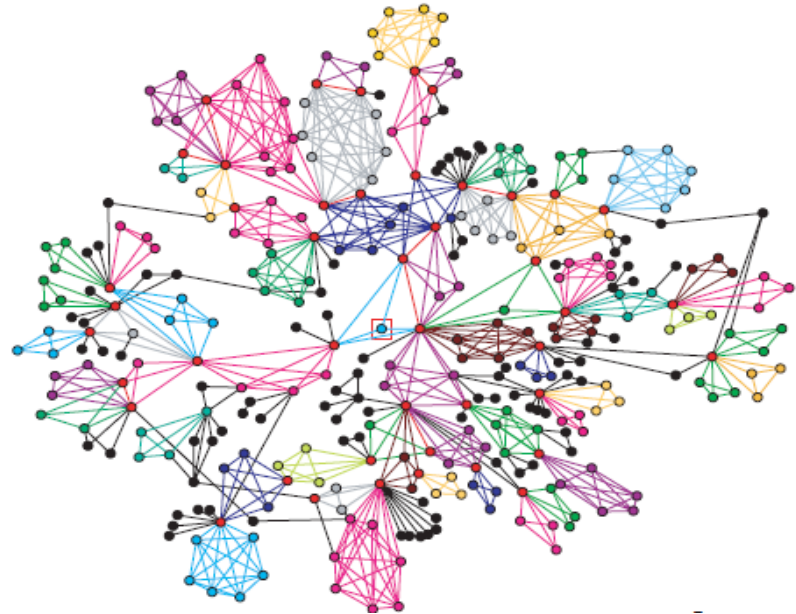
## 주요 네트워크 분석 기법 - Network structure

- Clustering coefficient
  - ✓ 나의 서로 다른 두 친구가 서로 친구일 확률

Low clustering coefficient



High clustering coefficient

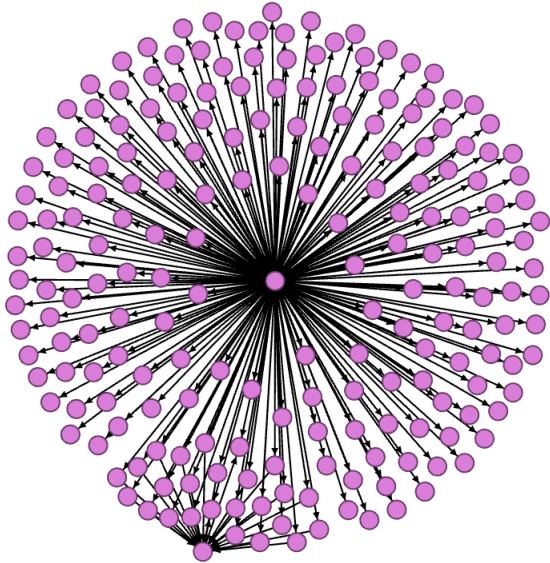


# Introduction

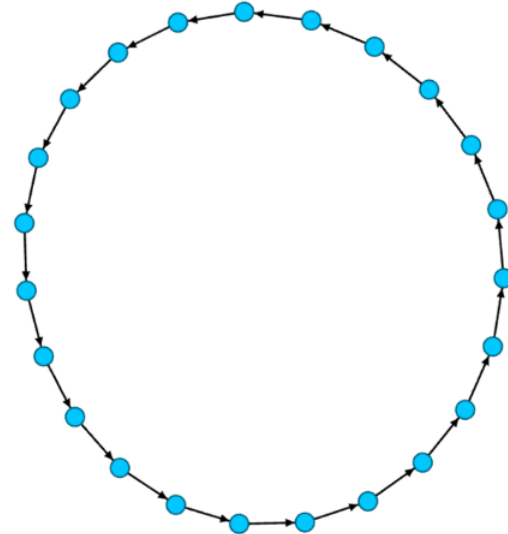
## 주요 네트워크 분석 기법 - Network structure

- Degree assortativity
  - ✓ 서로 연결된 노드쌍의 degree 상관관계

Low degree assortativity network



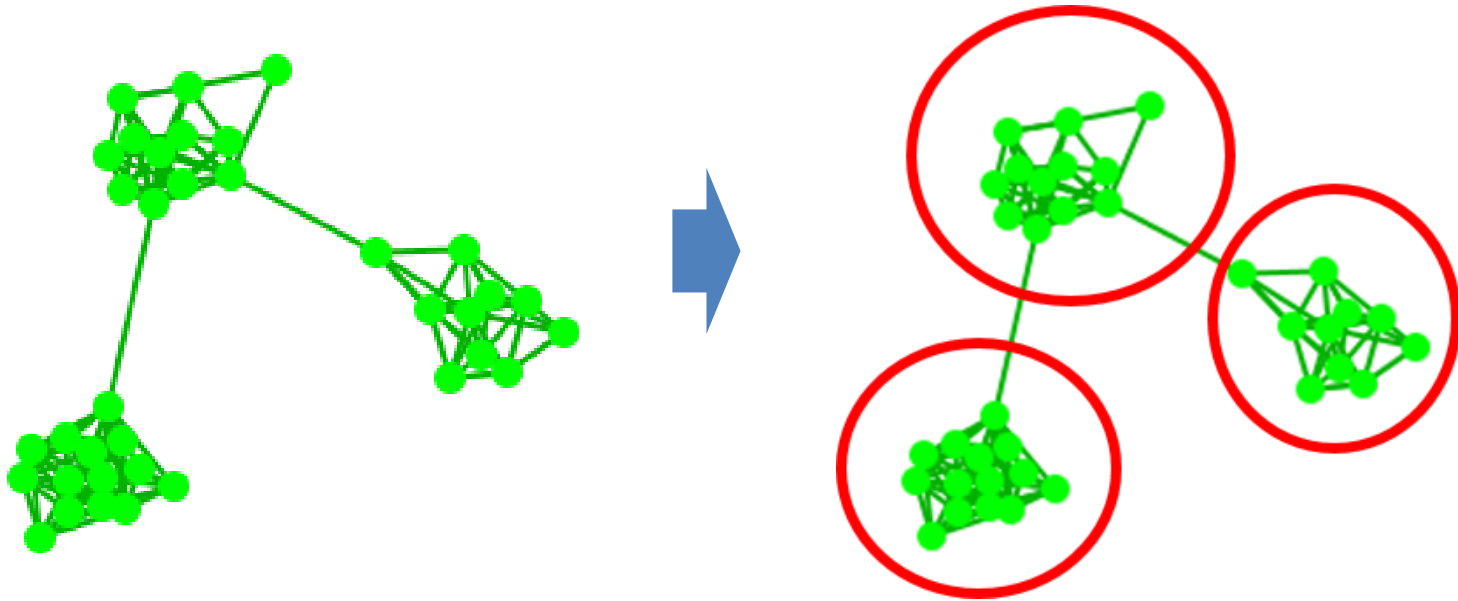
High degree assortativity network



# Introduction

## 주요 네트워크 분석 기법 - Community detection

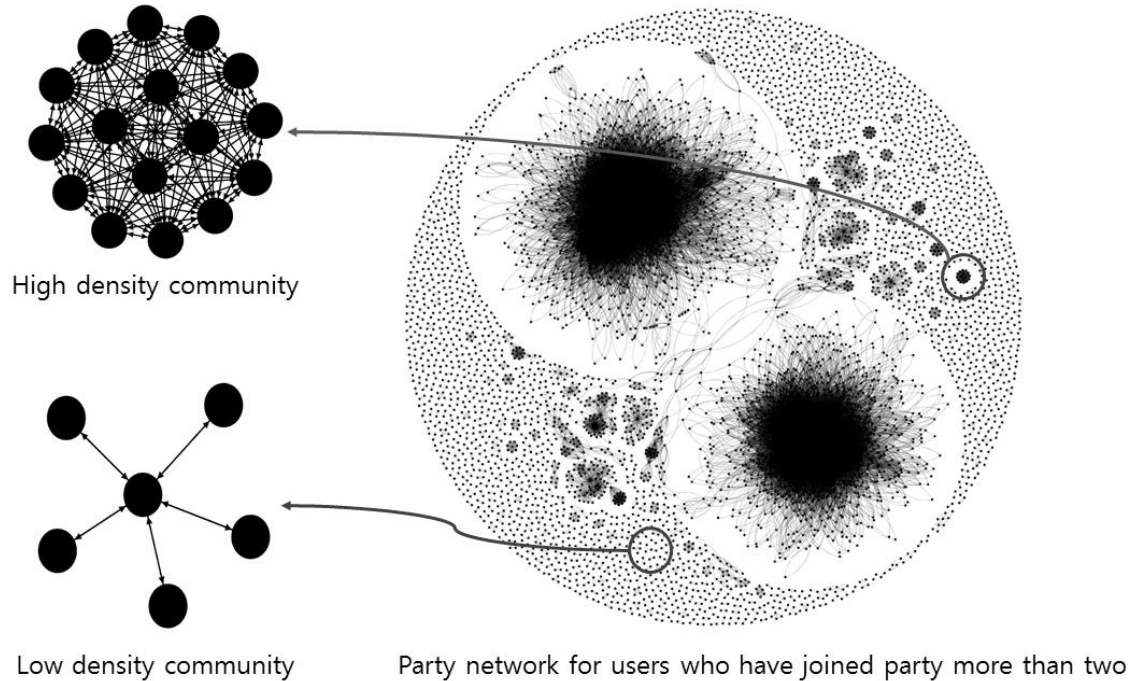
- 전체 네트워크를 연결 밀도가 상대적으로 높은 소집단끼리 분류



# Introduction

## 주요 네트워크 분석 기법 - Community detection

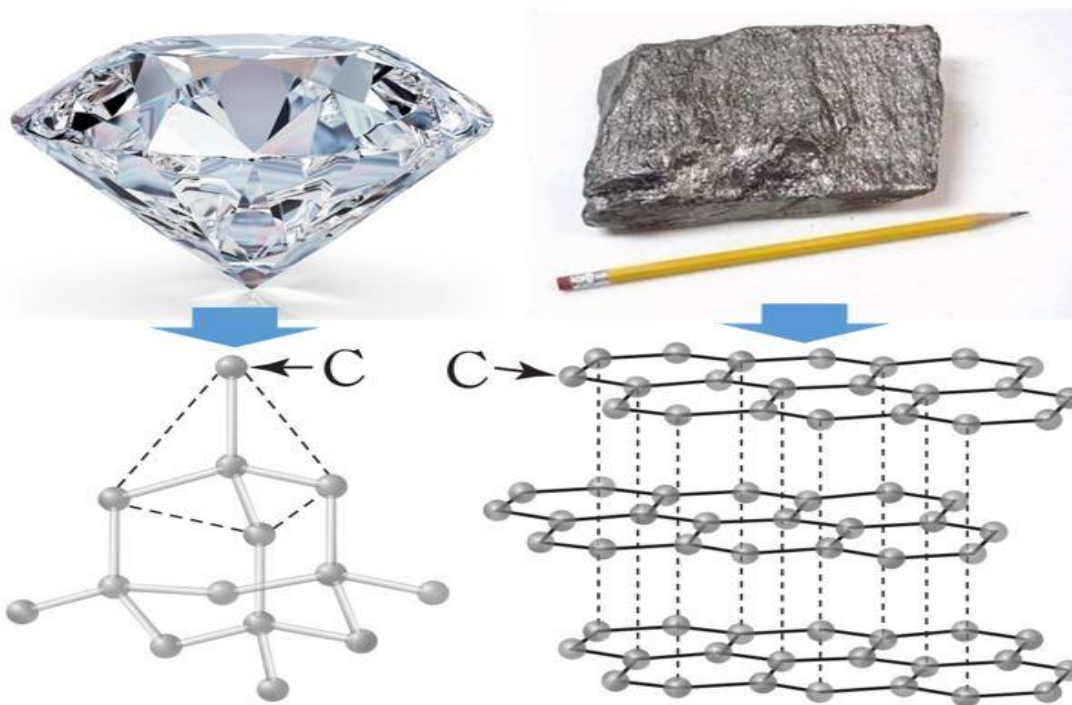
- 전체 네트워크를 연결 밀도가 상대적으로 높은 소집단 끼리 분류
  - ✓ 커뮤니티 간의 네트워크 구조 비교 분석



# Introduction

왜 네트워크 분석이 필요한가?

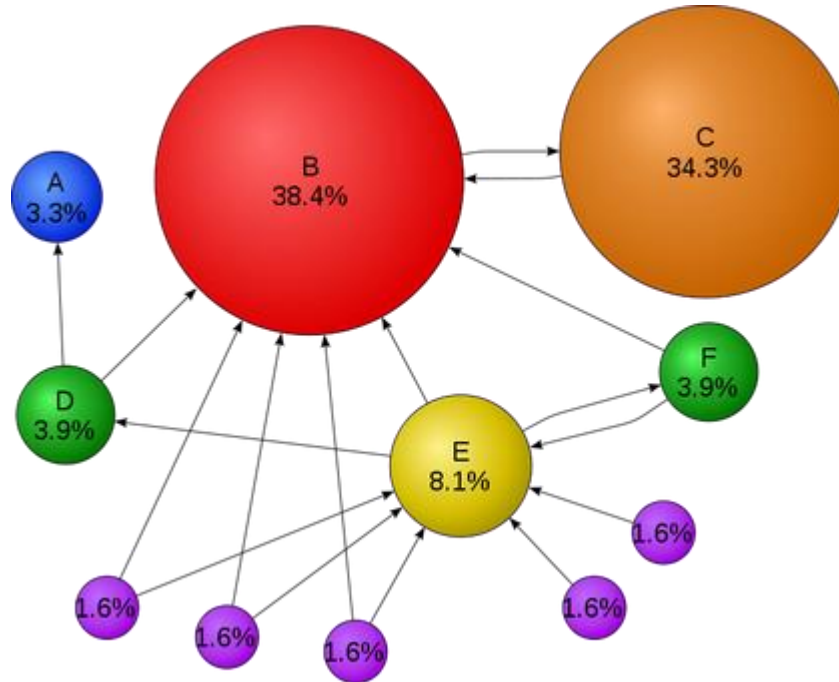
- 환원주의 오류
  - ✓ 다이아몬드와 흑연



# Introduction

## 왜 네트워크 분석이 필요한가?

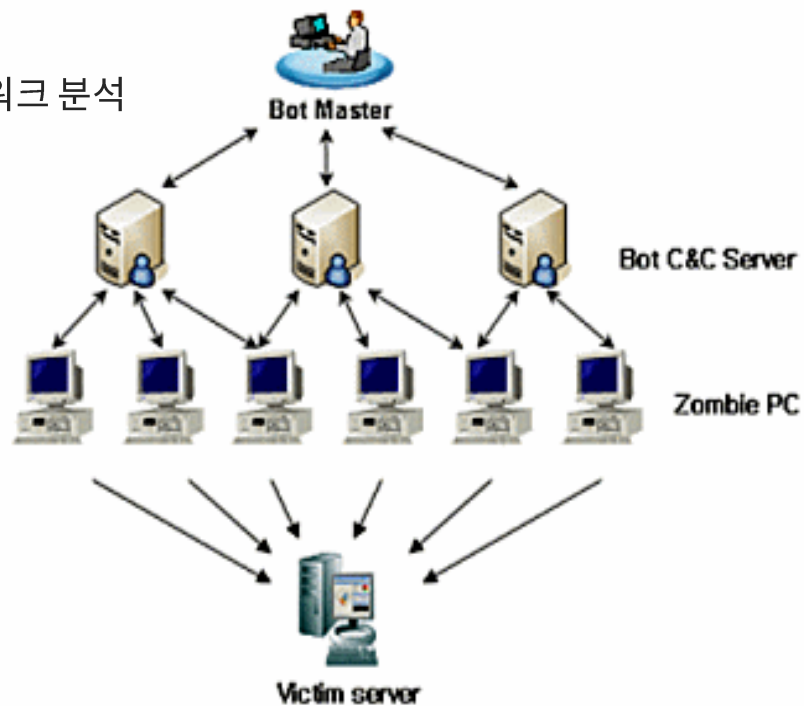
- 구글의 페이지 랭크
  - ✓ 웹 문서 자체가 가진 속성만으로 알 수 없는 문서의 중요도 측정



# Introduction

왜 네트워크 분석이 필요한가?

- Botnet 탐지
  - ✓ 악성 해킹을 통해 감염된 좀비 PC 들의 통신 네트워크 분석



# Introduction

## 온라인 게임에서는?

- 캐릭터 간의 다양한 상호 작용 및 관계 정보 발생
  - ✓ 거래, 파티, PvP, 채팅
  - ✓ 친구, 길드
- 유저의 특징을 정확히 분석하려면 네트워크 분석 중요
- 활용 사례
  - ✓ 유저 세그멘테이션
  - ✓ 이탈 예측
  - ✓ 악성 유저 탐지



# Introduction

## 사용 도구

- 분석 라이브러리
  - ✓ igraph
  - ✓ tidygraph
- 시각화
  - ✓ Gephi
  - ✓ ggraph

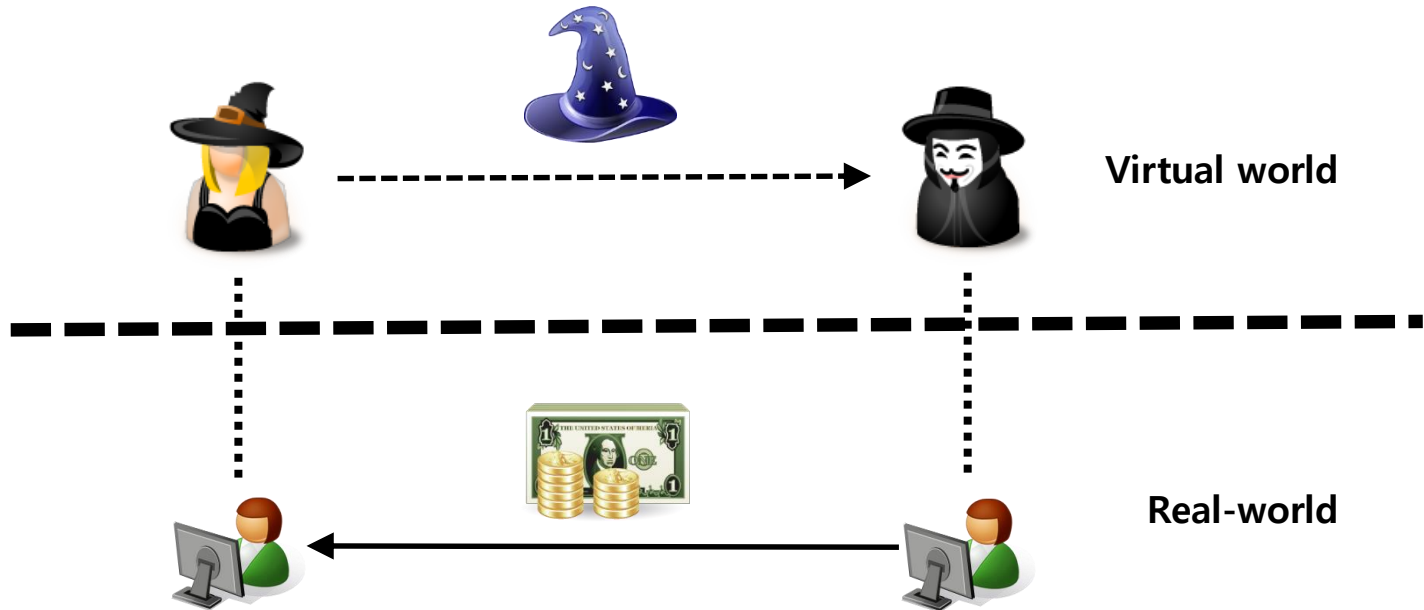
# 게임 데이터 분석 사례 #1

## 아이템 현금 거래량 추정

# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Real Money Trade (RMT)

- 게임 아이템을 현실 세계의 재화와 거래하는 행위



## 게임 데이터 분석 사례 #1 - 현금 거래 추정

### Yin and Yang of RMT

- Reasonable economic activity
- Key measurement for the popularity of an online game
- Contributions to a game's economic success and longevity



# 게임 데이터 분석 사례 #1 – 현금 거래 추정

## Yin and Yang of RMT

- Reasonable economic activity
- Key measurement for the popularity of an o
- Contributions to a game's economic succes



물품등록 리스트 25개 ▾

종류	등급	물품제목	단위금액	거래금액	등록시간
아데나	\$	[210만~10억] ★아데나 최저가 급처분★	70만당 6,880원	최소 20,640원	피워등록
아데나	G	[300만~9억5,700만] →안심/신속거래↔▶백만장7400원...	100만당 7,400원	최소 22,200원	17:14
아데나	G	[1억~9억5,700만] →안심/신속거래↔▶백만장7300원◀	100만당 7,300원	최소 730,000원	17:13
아데나	\$	[5,000만~1억6,000만] █████◆200%보상★손님은 왕█████	100만당 7,210원	최소 360,500원	17:13
아데나	V	[180만~2억] █████ 약값 소액판매 █████ 300%보상	20만당 1,960원	최소 17,640원	17:11
거래종료	B	[500] 1400만아데나 팝니다.	100만당 66,000,000원	33,000원	17:11
아데나	B	[180만~2억] █████ 약값 소액판매 █████ 300%보상	20만당 1,960원	최소 17,640원	17:10
거래종료	P	[400만] 남자답게 아덴 팝니다.	100만당 5,750원	23,000원	17:10
거래종료	V	[500만] 버투님	100만당 6,200원	31,000원	17:08
거래종료	V	[500만] 버투님	100만당 6,200원	31,000원	17:07
아데나	V	[40만] 아데나 팝니다. 24시간 풀대기중!	100만당 12,500원	5,000원	17:07
아데나	V	[20만] 아데나 팝니다. 24시간 풀대기중!	100만당 15,000원	3,000원	17:06
아데나	V	레벨보상 [180만~2억] █████ 약값 소액판매 █████ 300%보상	20만당 1,960원	최소 17,640원	17:05
거래종료	P	[400만] 교실	100만당 6,000원	24,000원	17:04
아데나	P	[300만~3,000만] █████◆안전*신용*믿음100%! ◆6단플...	100만당 7,900원	최소 23,700원	17:04
거래종료	P	[300만] 교실	100만당 6,000원	18,000원	17:02

## 게임 데이터 분석 사례 #1 - 현금 거래 추정

### Yin and Yang of RMT

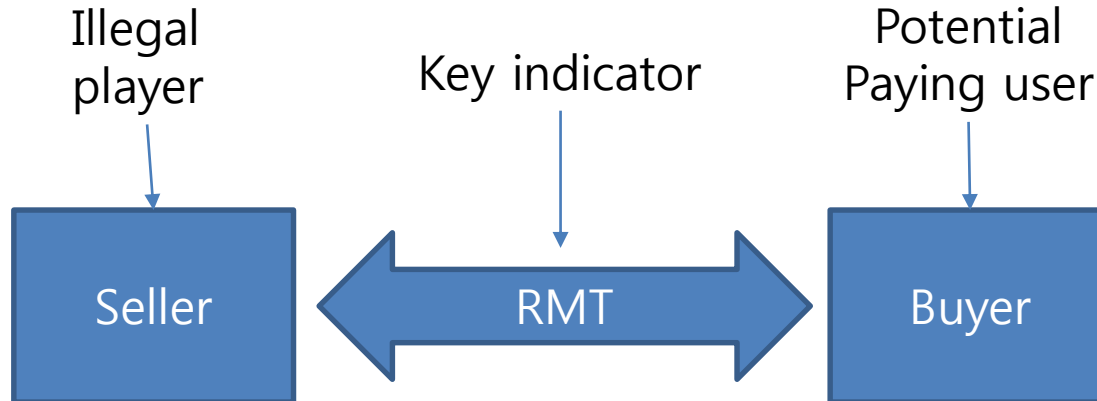
- **Gold Farming Groups (GFG)** involve in RMT



## 게임 데이터 분석 사례 #1 - 현금 거래 추정

Why is RMT analysis is so important in game industry

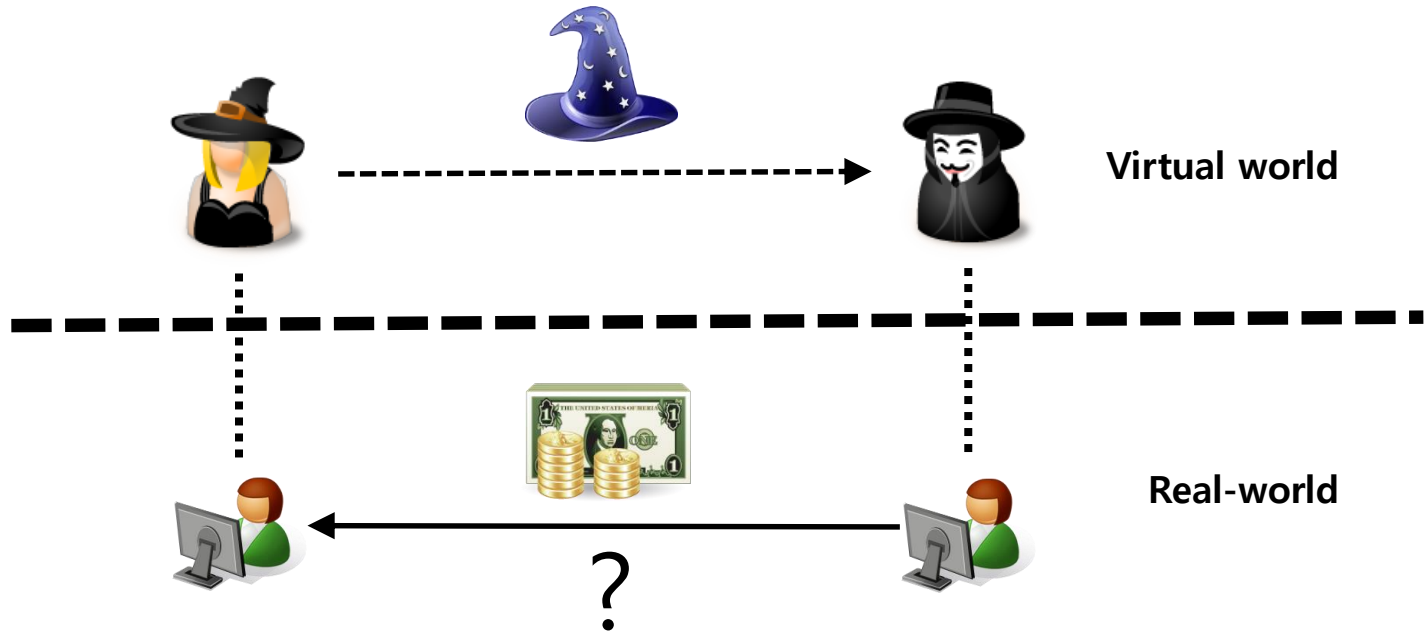
- RMT volume can be used as a **key measurement for the popularity** of an online game
- Tracking the transactions of RMT can **detect illegal users**
- Analyzing RMT consumers makes it possible to **find potential paying users**



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Challenges

- We can not trace currency flow in the real-world

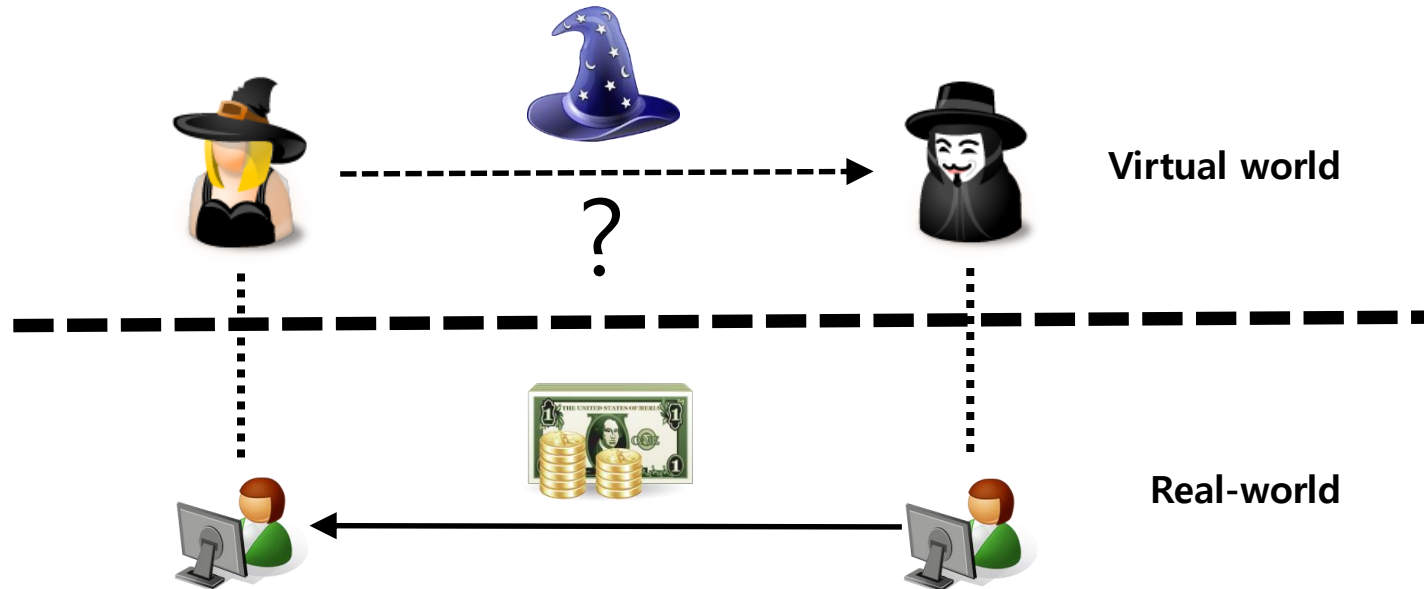




# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Challenges

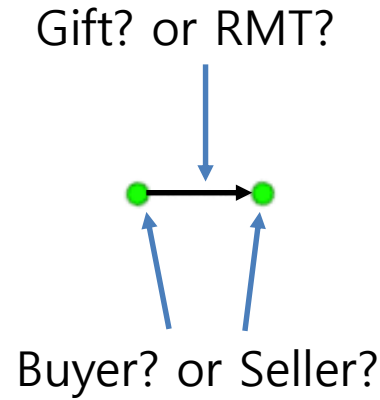
- In an on-line game, users conduct transactions for various purposes
- It is not easy to distinguish RMTs from normal item delivery transactions



## 게임 데이터 분석 사례 #1 - 현금 거래 추정

### Key idea

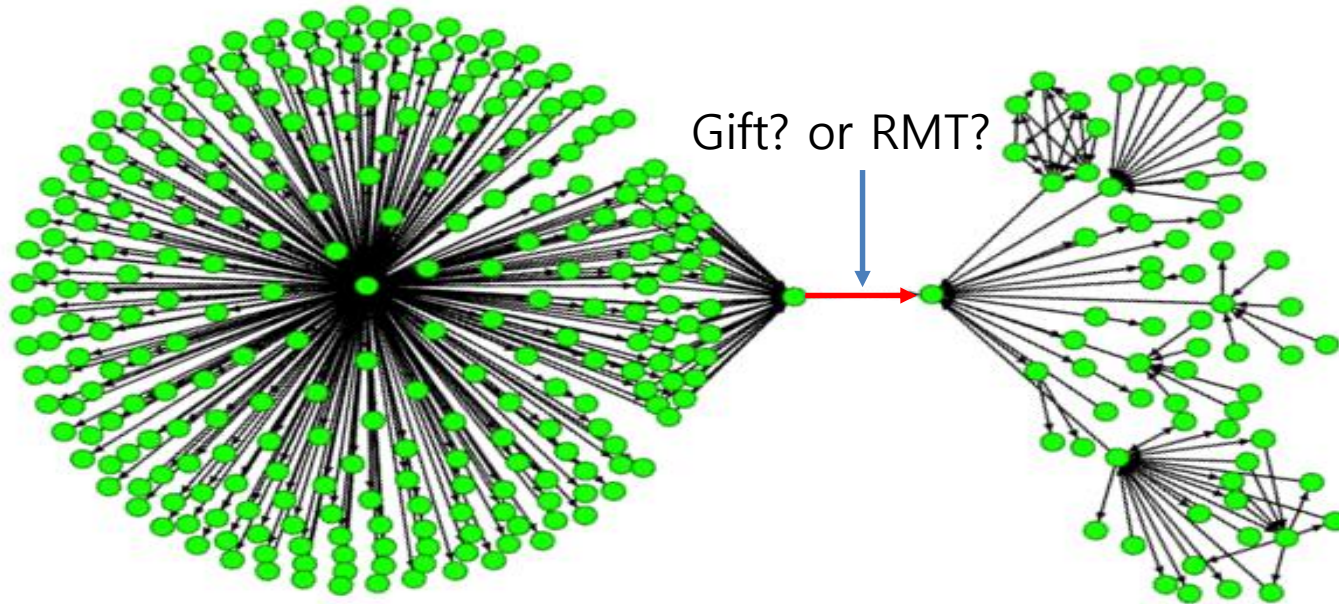
- Surely, it is not easy to identify RMT with a single transaction
- But, take a step back ...



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Key idea

- Surely, it is not easy to identify RMT with a single transaction
- But, take a step back **to look at the big picture**

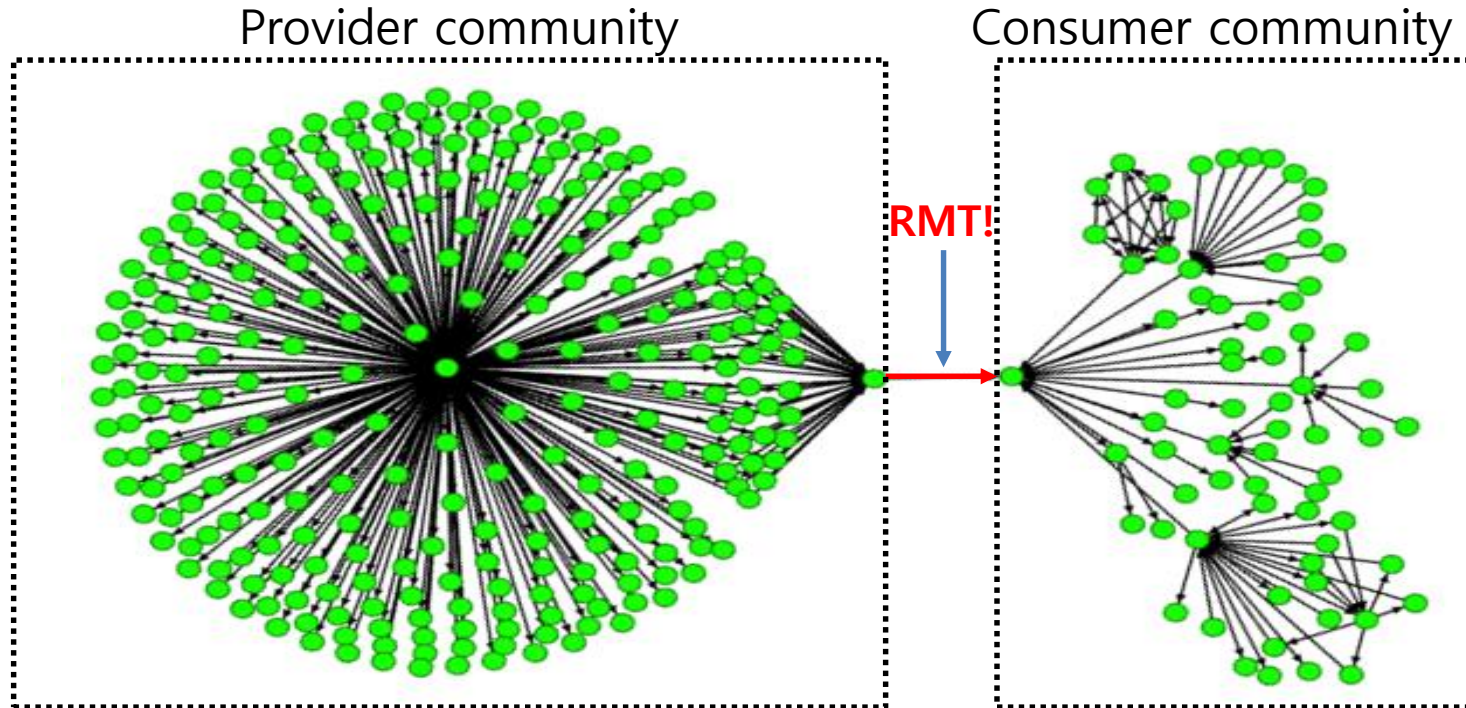


# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Key idea

➤ Focus on **trades between communities**, not users

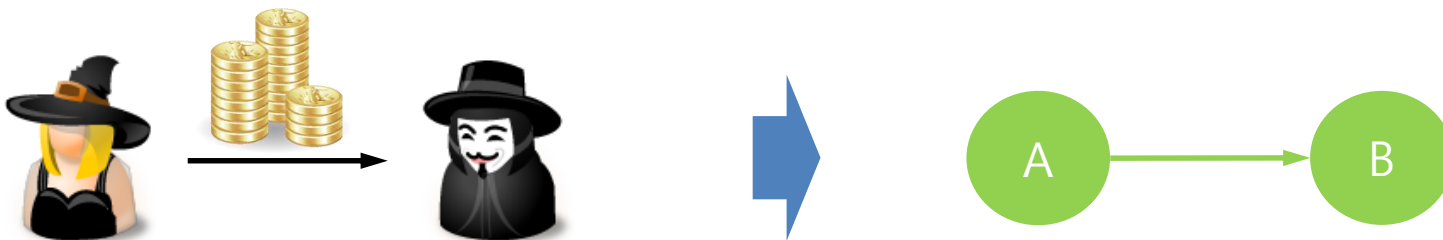
➤ Focus on identifying the **characteristics of communities**, not users



## 게임 데이터 분석 사례 #1 - 현금 거래 추정

### Construction of a virtual currency trading network

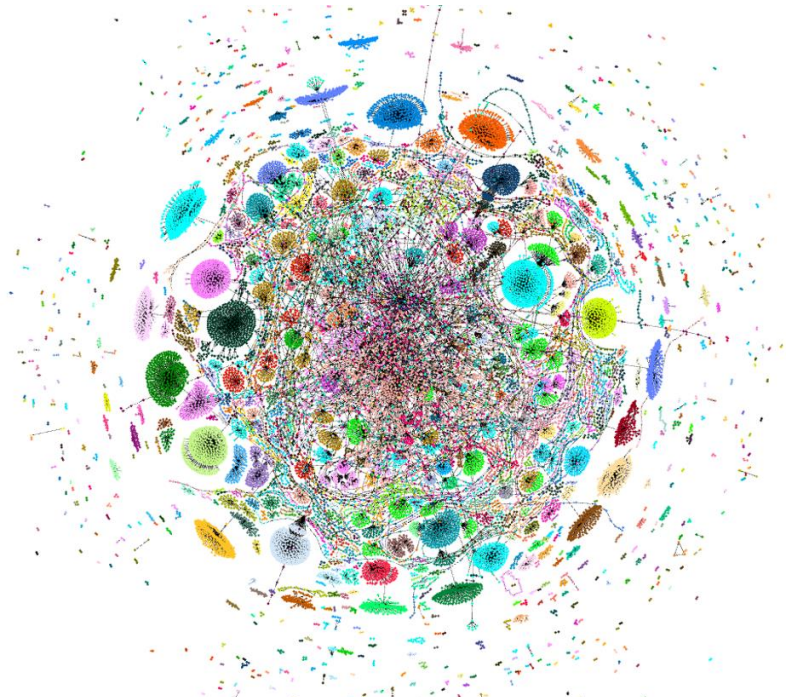
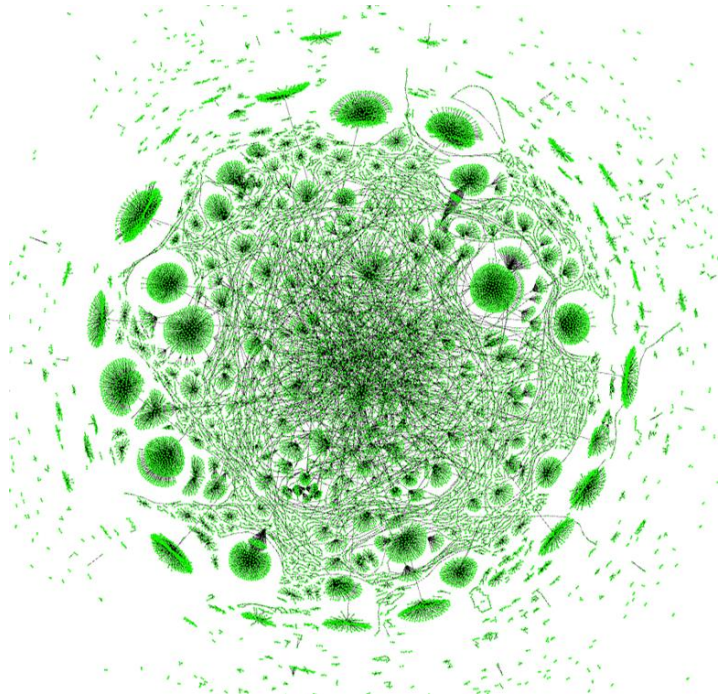
- Extracting only one-way trading logs and aggregating it on a weekly basis
- Constructing directed and weighted trading network
  - ✓ Nodes: game users
  - ✓ Edges: virtual currency trades between nodes
  - ✓ Weight: trade count



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Community detection

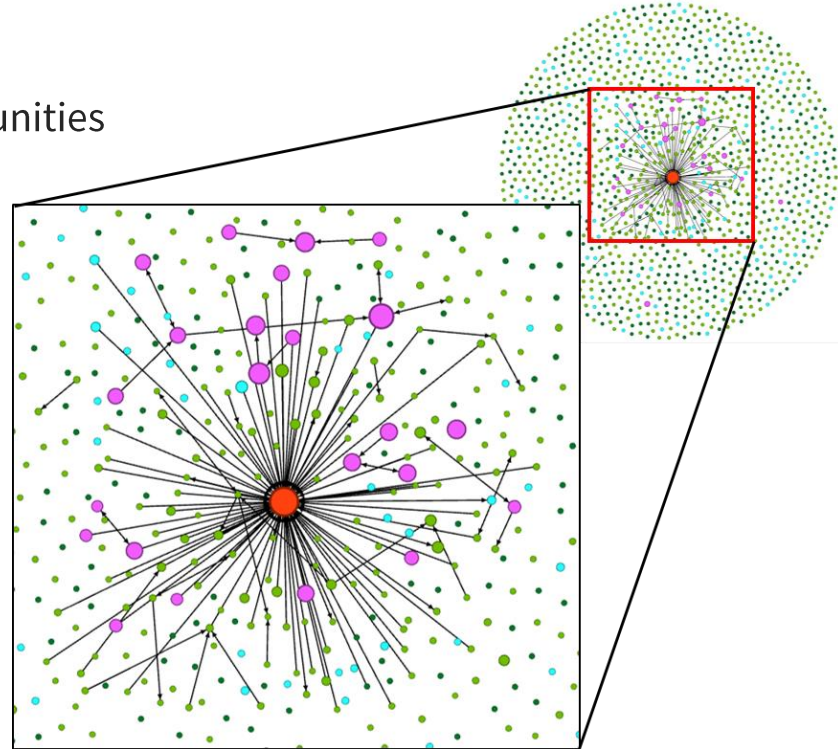
- Finding community structures in the constructed network



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Community detection

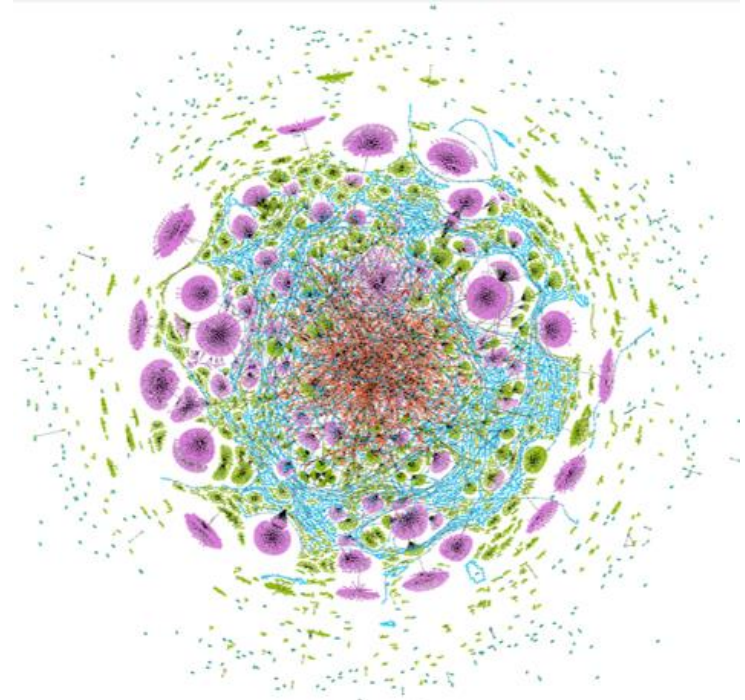
- Tracking transactions between communities



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Community network analysis

- Provider and consumer communities are detected by analyzing community characteristics
  - ✓ Network structures
  - ✓ In-game activities related with production and consumption of virtual currency
  - ✓ Payment amounts for the game company
  - ✓ Play style of members in communities



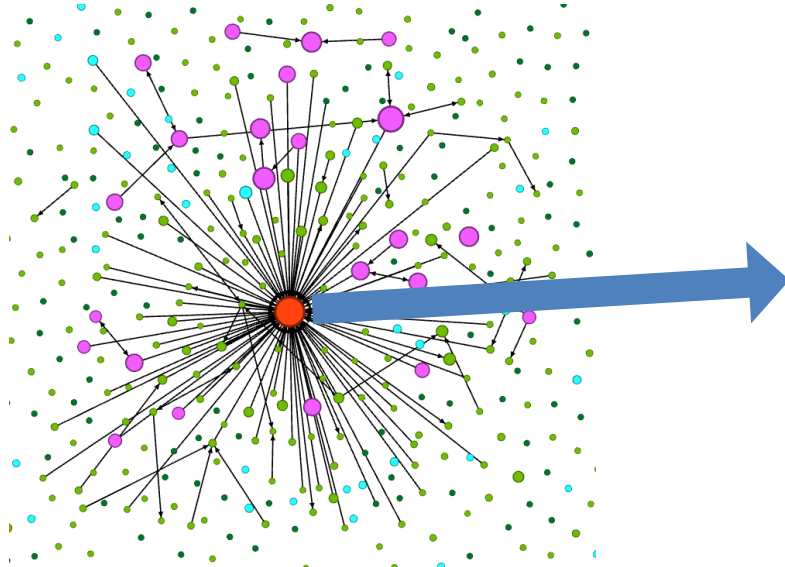


# 게임 데이터 분석 사례 #1 - 현금 거래 추정

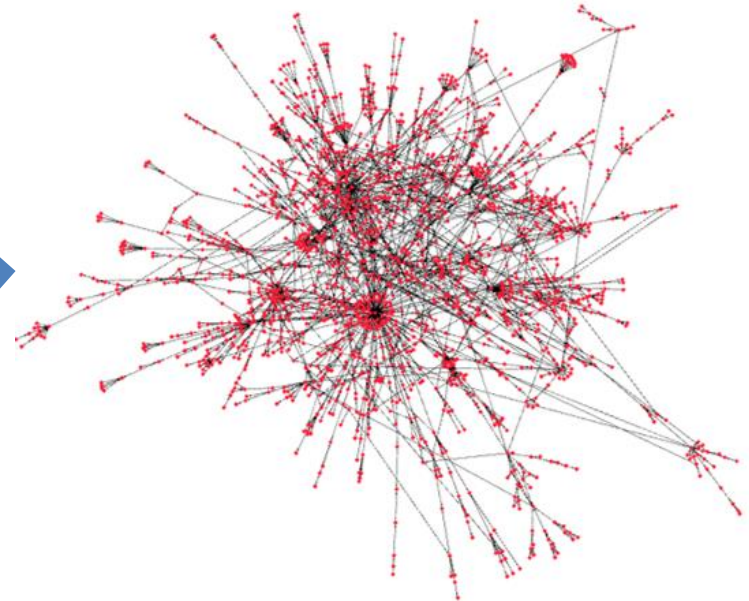
## Consumer community

- Absorbing virtual currency of inter-community trading
- **Scale-free network** like social media (Facebook, Twitter)

Community-based



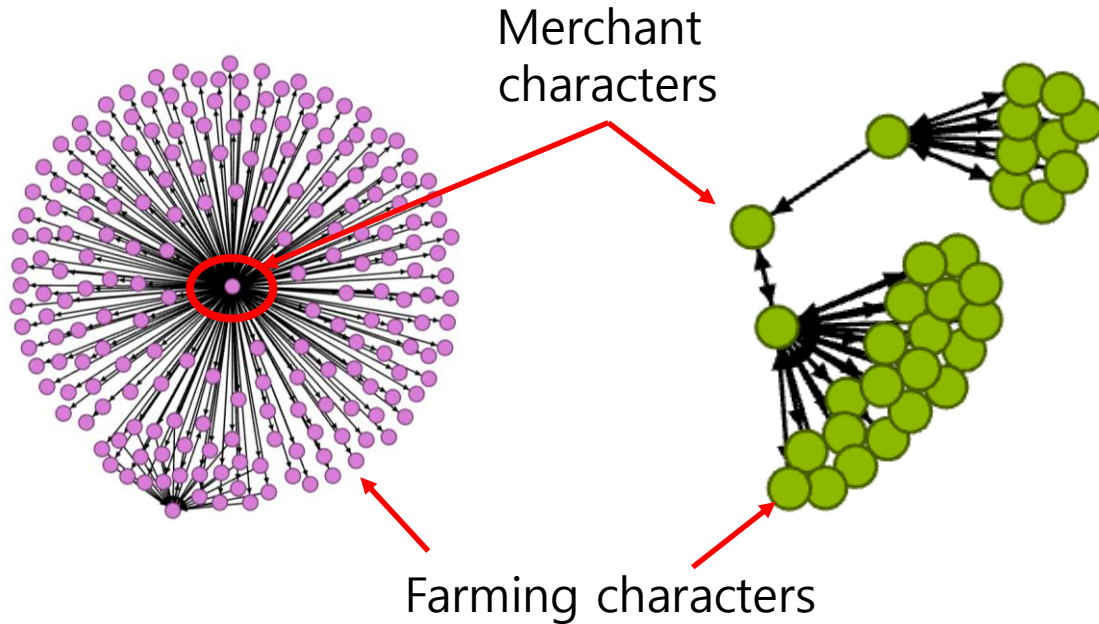
User-based



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Provider communities - Gold farming groups

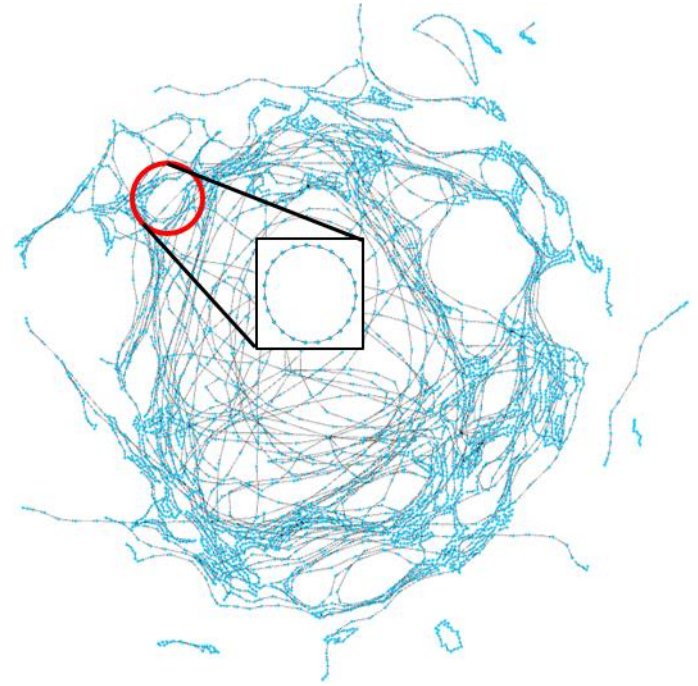
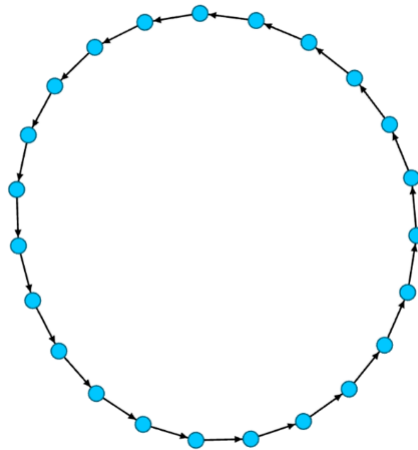
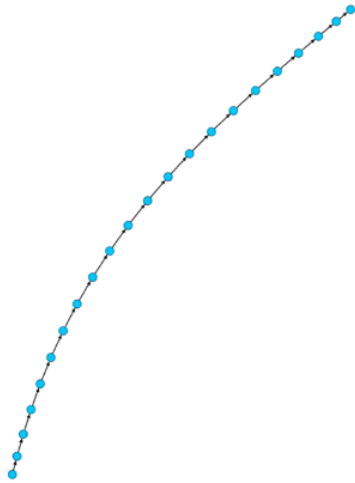
- Star-shaped network
- **Highly specialized organization** for efficiency



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Provider communities - Arbitrage groups

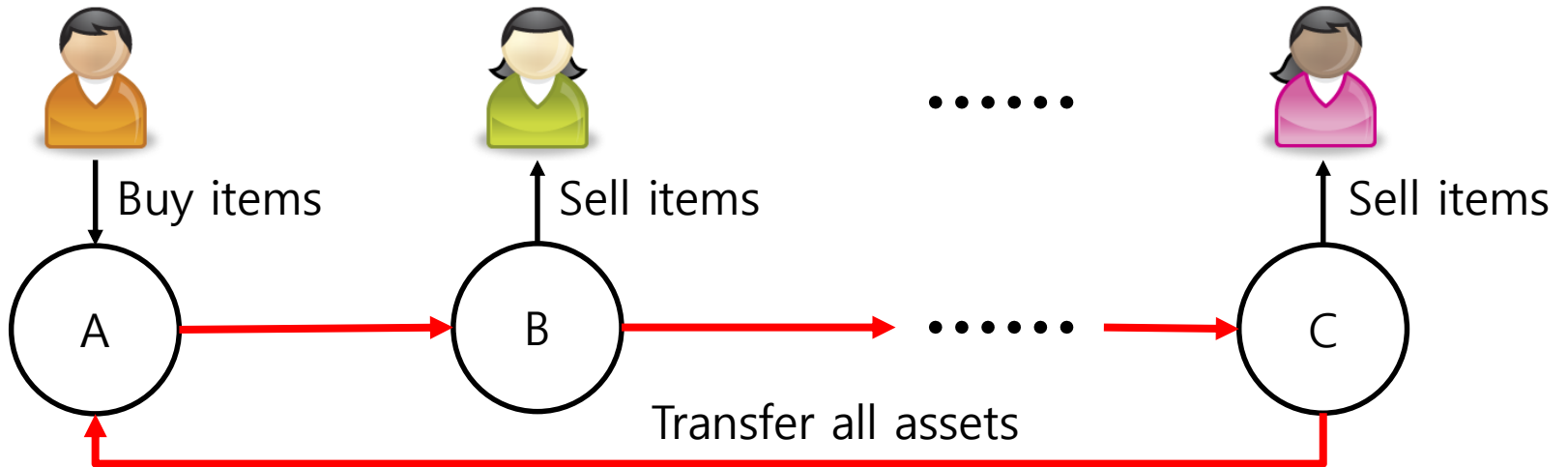
- One-way chain network
- **Extremely decentralizing** communities



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Provider communities - Arbitrage groups

- Possession of the largest amount of assets
- **Repeating buy and sell items with a profit margin** in the virtual world
- Using multiple accounts to evade in-game tax and suspicious transaction monitoring by game company

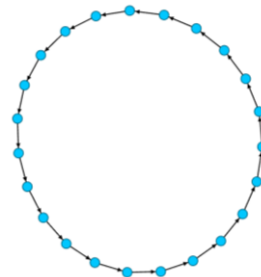
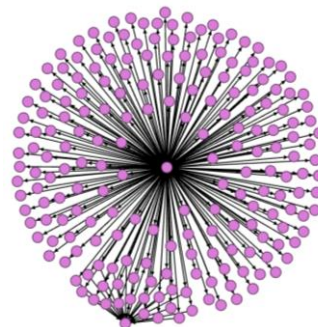
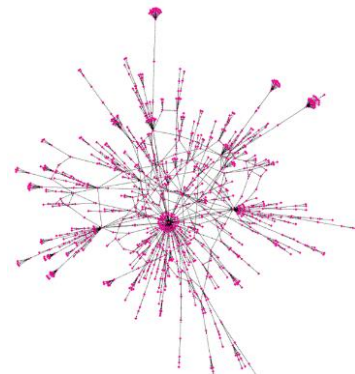
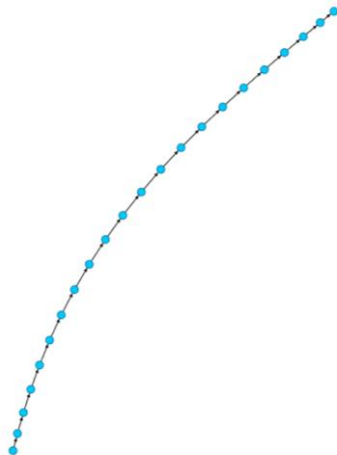
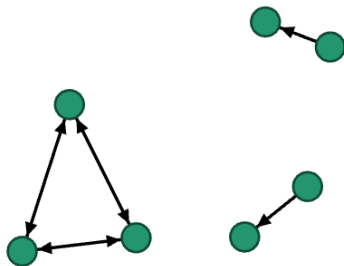


# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Community grouping by network structure

➤ Extracting the characteristics of community structure

- ✓ Mean / std. of degree
- ✓ Mean / std. of betweenness
- ✓ Degree assortativity
- ✓ Clustering coefficient
- ✓ Radius
- ✓ Community size



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Community grouping by network structure

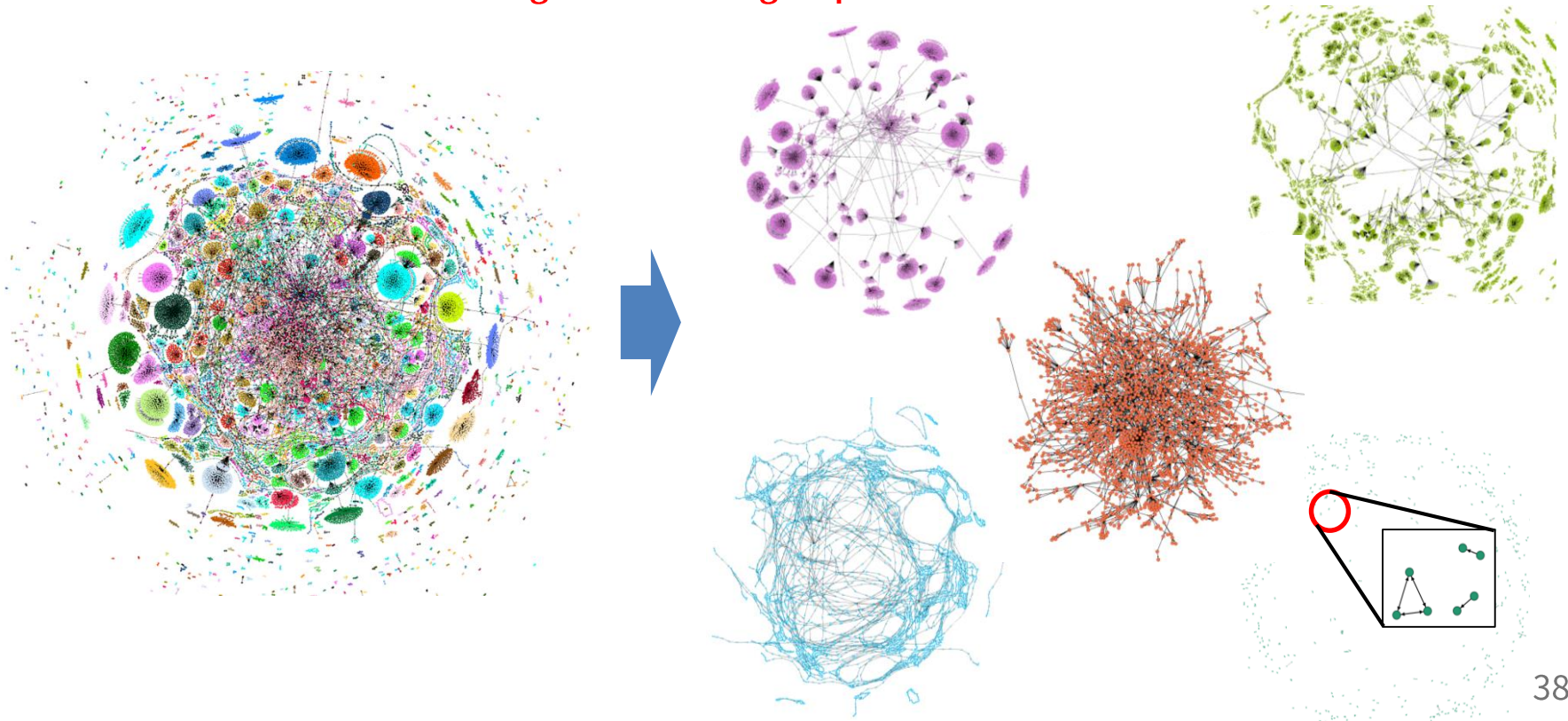
- Extracting the characteristics of community structure
  - ✓ Mean / std. of degree
  - ✓ Mean / std. of betweenness
  - ✓ Degree assortativity
  - ✓ Clustering coefficient
  - ✓ Radius
  - ✓ Community size

Community ID	Degree		Betweenness		Degree Assortativity	Radius	Size
	mean	std	mean	std			
1	1.3	0.57	0.3	0.57	-1	1	5
2	1.6	0.89	1.6	2.3	-0.67	1	14
3	2.25	14.4	107.4	1500.9	-0.83	2	213
4	2	0	117.2	30.9	0	16	32
5	2.29	4.43	3294	14163	-0.12	9	1033

# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Community grouping by network structure

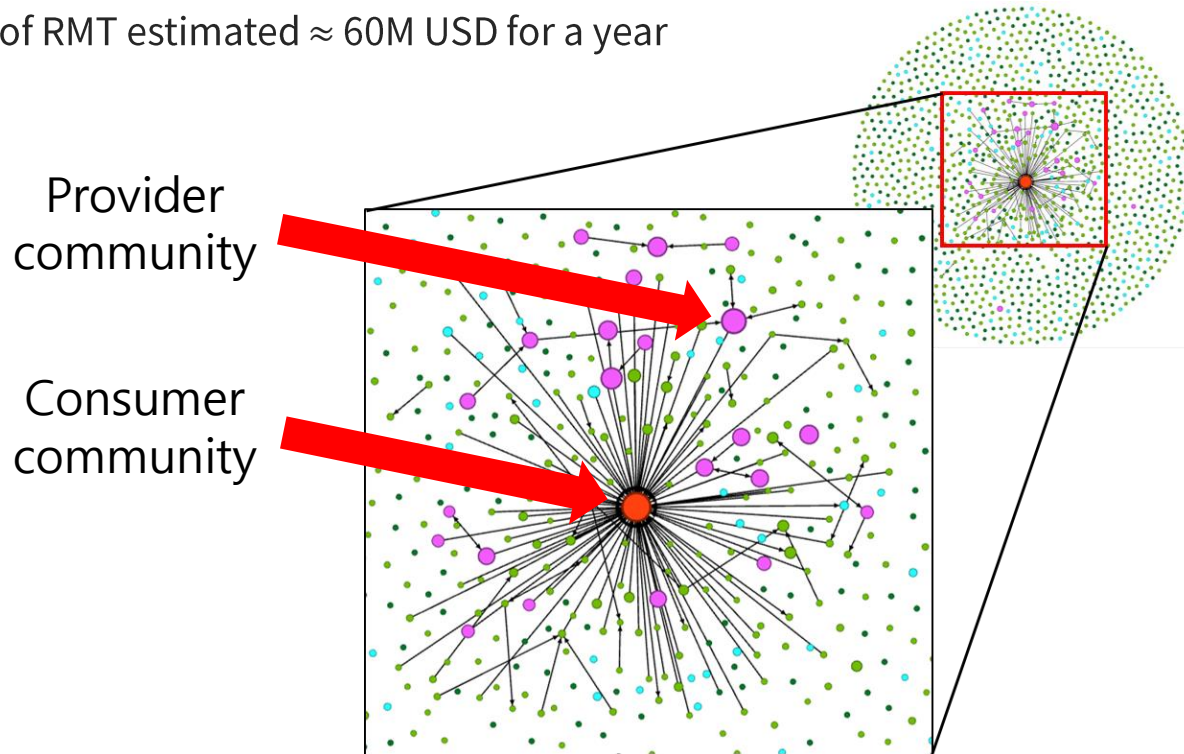
- The communities are **categorized into 5 groups**



# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## RMT estimation

- **Transactions between provider and consumer communities** are filter out as RMT
- Total volume of RMT estimated  $\approx$  60M USD for a year





# 게임 데이터 분석 사례 #1 - 현금 거래 추정

## Summary

- 게임 캐릭터 간에 어떤 재화를 전달할 때 그 이유를 파악하기 어려움
  - ✓ 부계정으로의 재화 이동
  - ✓ 친한 친구끼리 선물 혹은 같은 혈맹원에게 재화 분배
  - ✓ 아이템 현금 거래
  - ✓ ...
- 전체 거래 네트워크를 분석하여 거시적인 관점에서 맥락 파악
  - ✓ 전체 유저를 거래 네트워크 구조를 분석하여 여러 개의 소그룹으로 분류
  - ✓ 각 소그룹 (커뮤니티) 의 특징 분석을 통해 소비자 그룹과 공급자 그룹 탐지
  - ✓ 커뮤니티 단위의 거래 흐름 추적을 통해 전체 RMT 규모 추정

# 게임 데이터 분석 사례 #2

## 유저간 전투 분석

## 게임 데이터 분석 사례 #2 - 유저간 전투 분석

### 배경

- MMORPG 에서 유저간 전투가 차지하는 비중은 매우 높음
- 전투가 게임의 몰입도 및 서비스 수명에 미치는 영향 높음
- 전투에 영향을 주는 요인 및 전투가 유저에게 미치는 영향에 대한 체계적 분석 필요



## 게임 데이터 분석 사례 #2 - 유저간 전투 분석

### 분석의 어려움

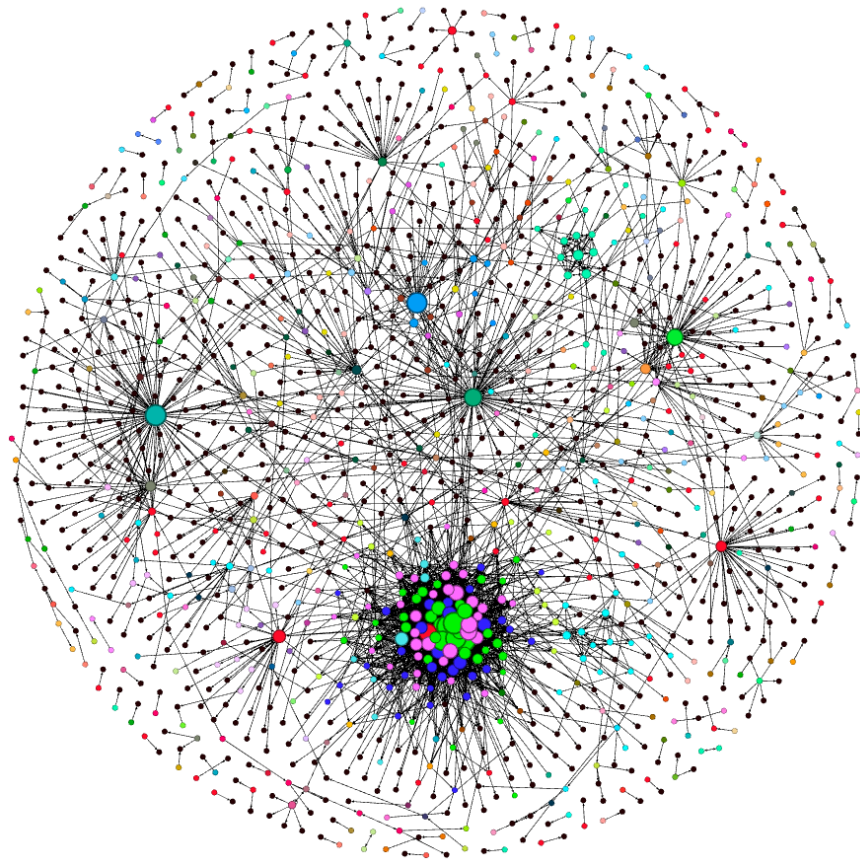
- 유저간 전투 목적은 다양함
  - ✓ 일방적인 학살 (막피)
  - ✓ 혈맹 간의 전투
  - ✓ 우연한 다툼
  - ✓ 친밀감의 표시



## 게임 데이터 분석 사례 #2 - 유저간 전투 분석

### 전투 네트워크 분석

- 캐릭터 간의 전투를 네트워크로 구성
  - ✓ 노드: 캐릭터
  - ✓ 엣지: 전투 행위 (방향: 선공)
- 네트워크 구조를 통해 전투 유형 분류
  - ✓ 단체전 (혈맹전)
  - ✓ 막피 (일방적 학살)
  - ✓ 단발성 전투

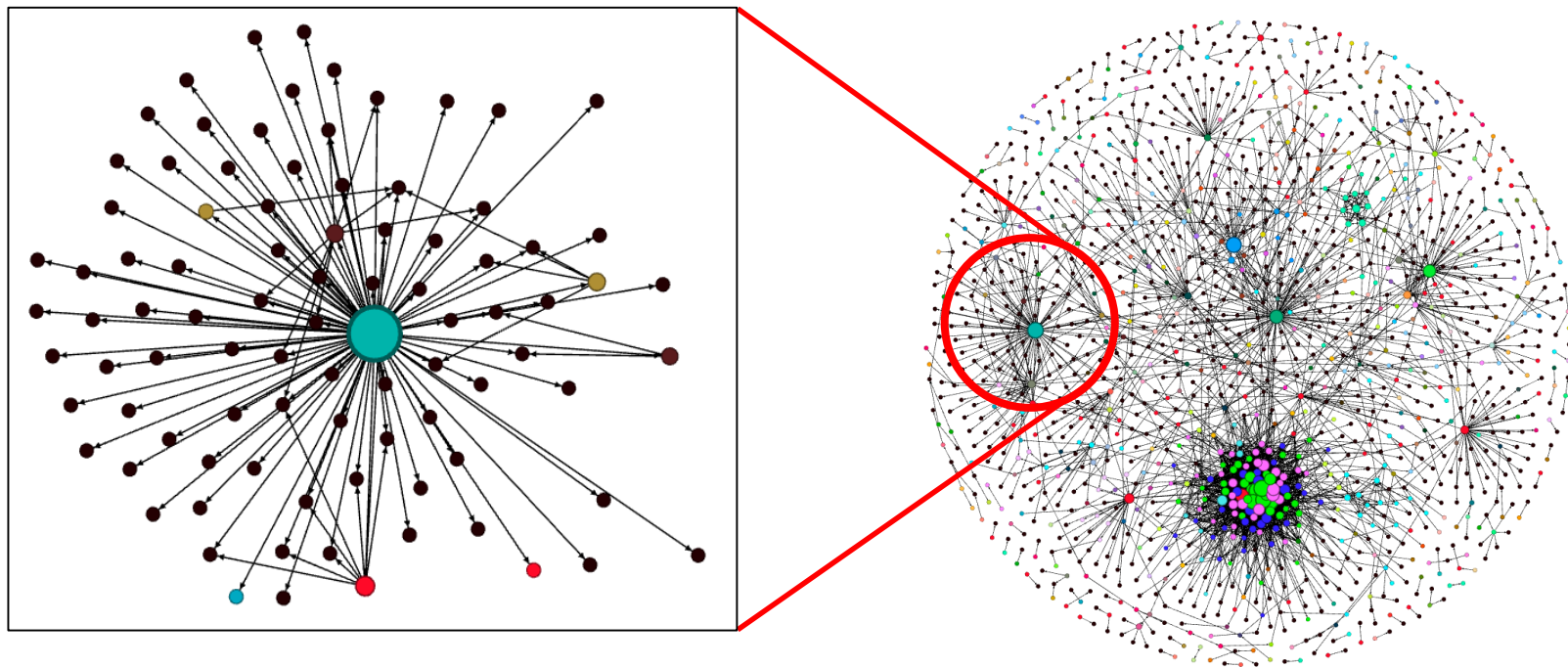


## 게임 데이터 분석 사례 #2 - 유저간 전투 분석

### 전투 네트워크 분석

➤ 막피 전투

- ✓ 하나의 캐릭터가 다수의 캐릭터와 전투

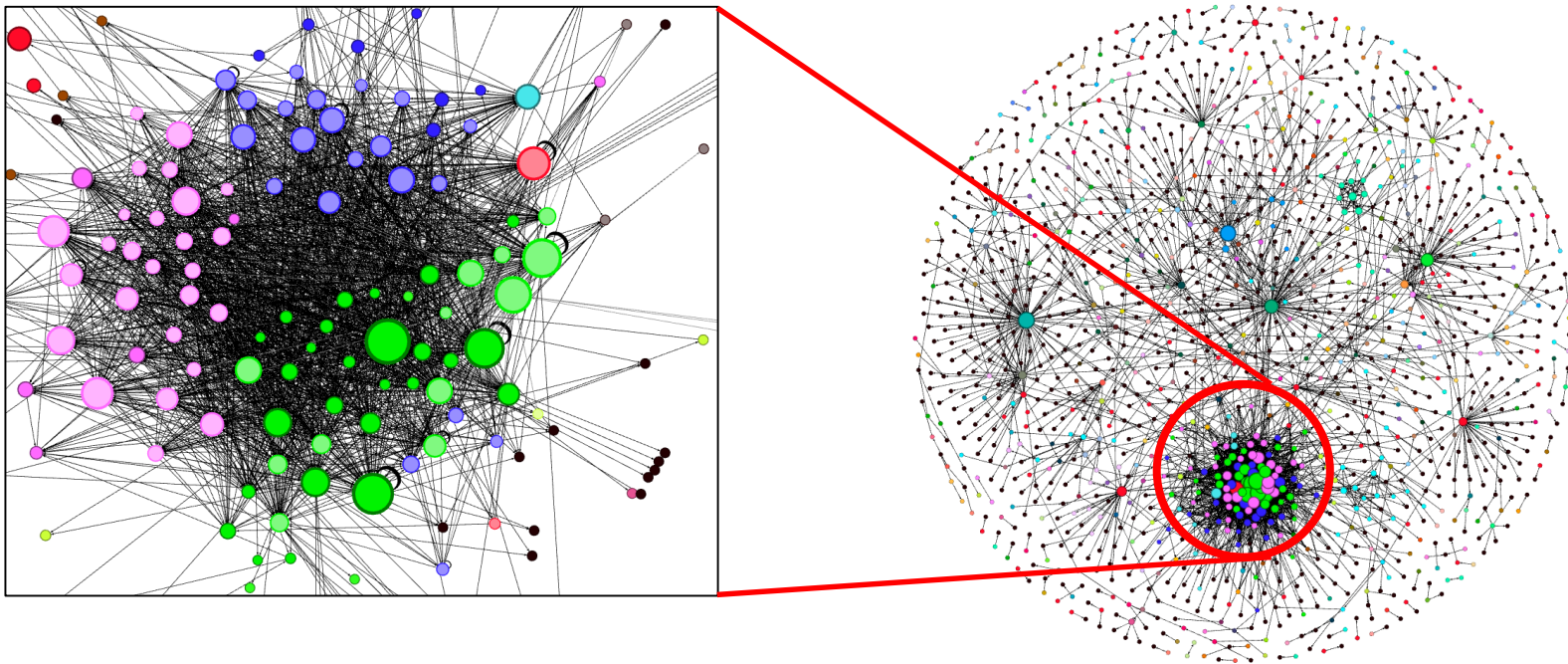


# 게임 데이터 분석 사례 #2 - 유저간 전투 분석

## 전투 네트워크 분석

### ➤ 혈맹전

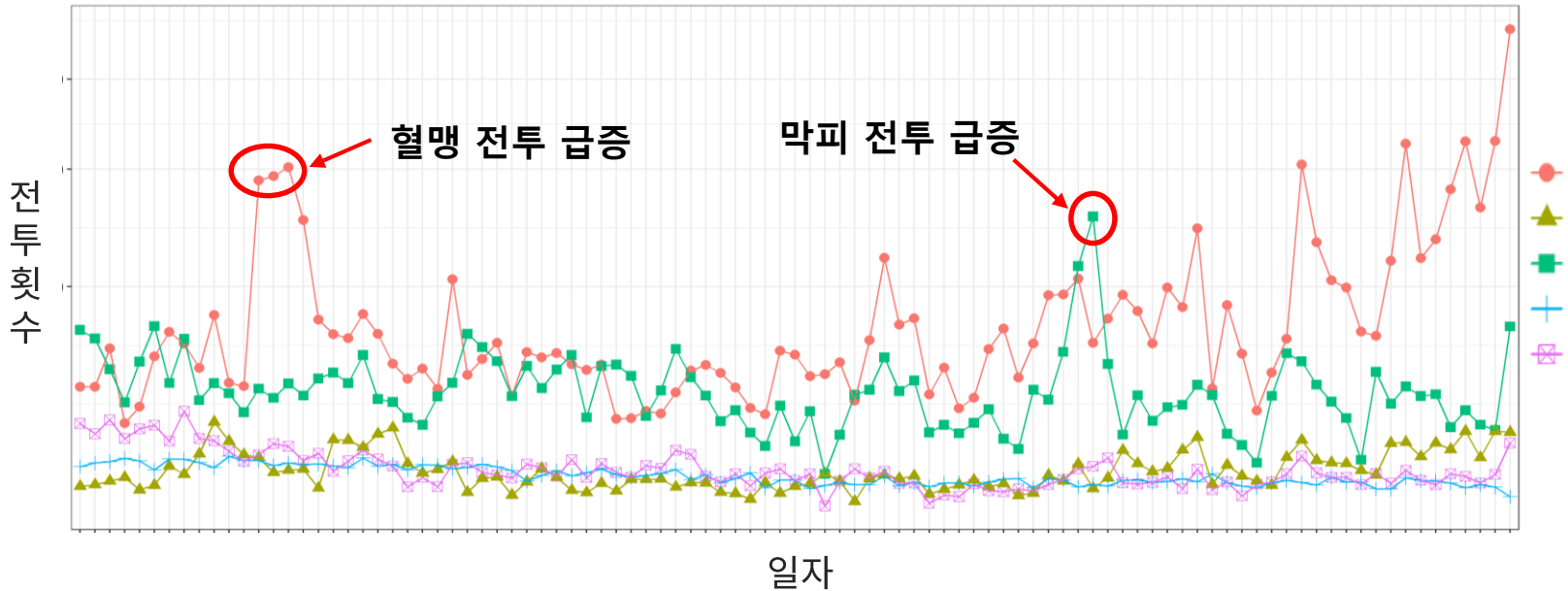
- ✓ 둘 이상의 혈맹에 속한 다수의 캐릭터 간의 분쟁



# 게임 데이터 분석 사례 #2 - 유저간 전투 분석

## 전투 유형 분류

- 개별 전투 유형 분류
  - ✓ 요인별 전투 동향 모니터링 및 대응

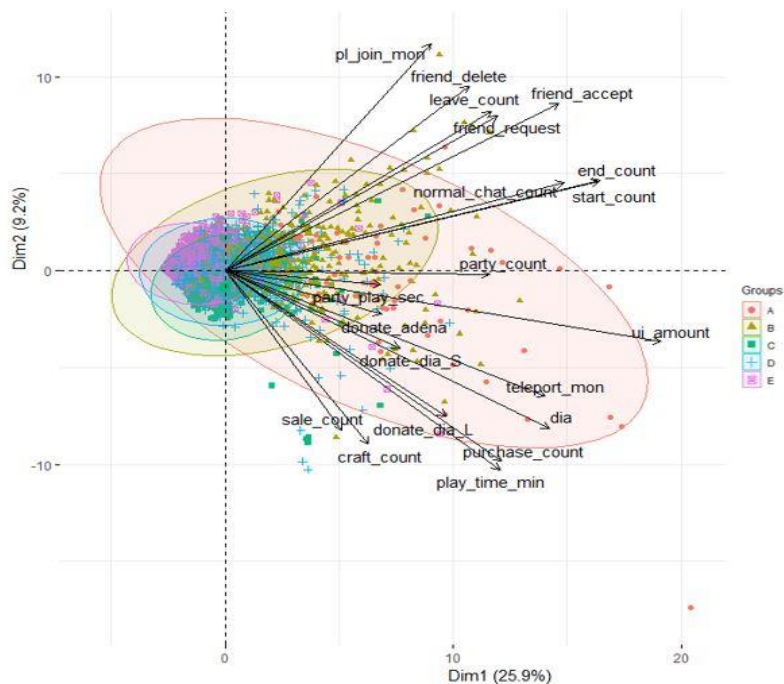
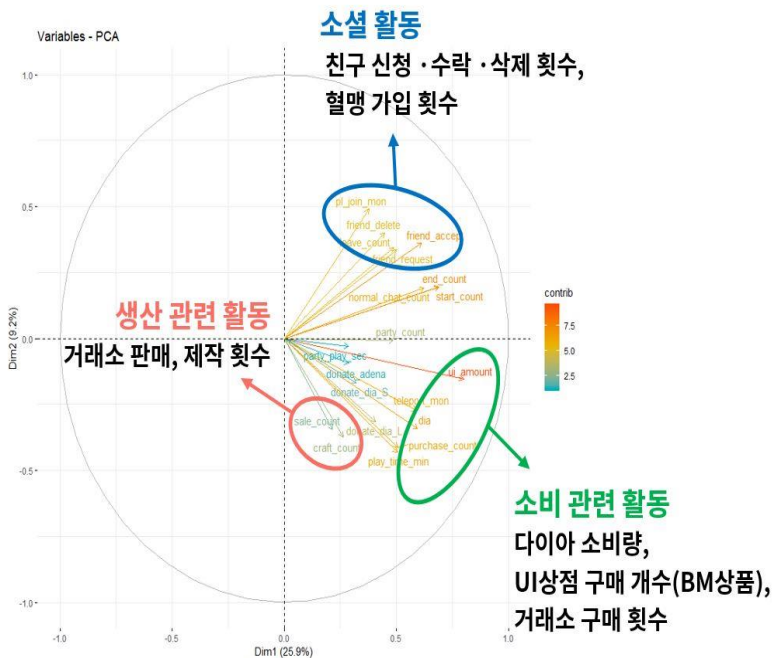




# 게임 데이터 분석 사례 #2 - 유저간 전투 분석

## 전투 유형 분류

- 캐릭터 별 전투 유형 분류
  - ✓ 선호하는 전투 유형에 따른 유저의 특징 관찰



# 결론

# 결론

## Summary

- 네트워크 분석의 장점
  - ✓ 개체 간의 관계 구조를 통해 거시적인 특성 파악 가능
  - ✓ 개체에 대한 **환원주의적 분석을 보완**해 줄 수 있음
- 네트워크 분석 기법
  - ✓ Node centrality
    - Degree, betweenness, closeness, ...
  - ✓ Network structure
    - Density, radius, clustering coefficient, degree assortativity, ...
  - ✓ Community detection
- 게임 데이터 분석: 유저 간의 다양한 사회 활동에 대한 분석 가능
  - ✓ 거래 네트워크
  - ✓ 전투 네트워크

# 결론

## 향후 발전 방향

- Representation learning for network structure
  - ✓ Node2vec
    - Centrality
    - Community detection
  - ✓ Graph Convolutional Networks
    - Network structure

**Q&A**