

병원 간 환자 이동 네트워크 시각화

곽 규 한¹ · 박 진 우² · 송 현 주^{3*}¹송실대학교 컴퓨터학부 석사과정²송실대학교 컴퓨터학부 학사과정^{3*}송실대학교 컴퓨터학부 교수

Inter-Hospital Patient Transfer Network Visualization

Kyuhan Kwak¹ · Jinu Park² · Hyunjoo Song^{3*}¹Master's Course, School of Computer Science and Engineering, Soongsil University, Seoul 06978, Korea²Bachelor's Course, School of Computer Science and Engineering, Soongsil University, Seoul 06978, Korea^{3*}Professor, School of Computer Science and Engineering, Soongsil University, Seoul 06978, Korea

[요 약]

의료 기관 간의 효과적인 연계의 중요성이 다양한 질병에 대한 치료 과정에서 부각되고 있다. 예를 들어 뇌졸중의 경우 EVT (Endovascular thrombectomy)의 도입 이후, 효과적인 치료 예후에도 불구하고 도입 금액이나 인적 자원 등으로 인하여 모든 병원에 도입하기 어렵다는 한계가 있으며, 이와 관련하여 병원 간의 연계 상황에 대한 연구나, 네트워크 형성 및 운영에 대한 연구가 진행되고 있다. 여러 선행 연구에서는 이를 위하여 병원 간의 환자 이동 정보를 기반으로 한 네트워크를 도출하고, 해당 네트워크에 대한 정량적 분석과 함께 시각적인 분석을 다양하게 시도해왔다. 하지만 대부분의 선행 연구들은 노드-링크 다이어그램의 형태를 차용한 시각화를 활용했다. 이러한 시각화 방법은 병원의 수가 늘어남에 따라 시각적 혼란을 야기할 수 있으며, 더 나아가 다른 형태의 시각적 표현을 통해 기존보다 더 많은 정보를 표현할 수 있다. 본 논문에서는 네트워크 시각화에 대한 선행 연구들을 기반으로 데이터의 형태와 분석 목적에 따른 효과적인 시각화 방법을 제안한다.

[Abstract]

The importance of effective coordination between medical institutions is emerging in the treatment process for various diseases. For example, EVT (i.e., Endovascular thrombectomy) raised the demand for close coordination between medical centers for stroke patients. Despite the effective treatment prognosis, it is challenging to introduce EVT to all hospitals due to the cost and personnel limitations. To this end, several previous studies have inferred a network based on patient transfer data between hospitals and attempted visual analysis along with quantitative analysis of the network. However, the majority of prior work relied on variants of node-link diagrams for visual analysis. As the number of hospitals grows, such visualization could suffer from visual clutter, and the data could reveal additional information with other visual representations. In this paper, we propose effective visualization methods according to the data type and analysis purpose based on previous studies on network visualization.

색인어 : 환자 이동, 병원 간 네트워크, 네트워크 시각화, 커뮤니티 시각화, 시각적 분석**Keyword** : Patient Transfer, Inter-hospital Network, Network Visualization, Community Visualization, Visual Analysis<http://dx.doi.org/10.9728/dcs.2023.24.2.411>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 16 December 2022; Revised 10 January 2023

Accepted 10 January 2023

***Corresponding Author; Hyunjoo Song**

Tel: +82-2-820-0676

E-mail: hsong@ssu.ac.kr

I. 서 론

의료 기관 간의 효과적인 연계에 대한 연구가 뇌졸중의 치료 과정과 치료 결과의 개선에 있어서 중요한 역할을 하고 있다 [1]. 특히 뇌졸중의 경우 endovascular thrombectomy (EVT)의 도입 과정에서 의료 기관 간의 연계의 효용과 중요성이 대두되었다 [2]. EVT의 효과는 2015년에 진행된 여러 연구[3], [4]에서 입증되었으나, 모든 병원에서 이를 도입하기에는 여건상 어려웠다. 이에 따라 일부 주요 병원들에 환경을 갖추고 주변 병원에서는 환자가 발생했을 때 EVT가 가능한 병원으로 환자를 이송하는 형태를 갖추게 되었다. 그 결과 뇌졸중 환자의 이송에 대해서 관심을 갖고 어떤 형태로 병원 간의 네트워크가 구성되는지에 대해 연구가 진행되고 있다 [1], [5].

한국에서는 병원 간의 연계를 위하여 상급종합병원을 중심으로 한 17개의 권역과 함께 70개의 중진료권을 설정하여 관리하고 있다 [6]. 하지만 정부에서 설정한 지역적 범위와 의료 기관들이 상호 경쟁에 의해 시장을 점유하는 지역적 범위 사이에는 차이가 있을 수 있다. 이로 인해 진료권 설정 방법에 대해서는 아직도 논의가 계속되고 있는 상황이며, 실제로 어떻게 병원 간 연계가 되고 있는지 확인하고자 노력하고 있다. 그러나 기존 연구에서 언급된 것처럼 뇌졸중 치료를 위한 네트워크 구성이 어떻게 되어 있고, 전 권역에 대해서 효과적으로 동작하고 있는지를 확인할 수 있는 정보는 아직 알려져 있지 않은 상태이다 [7].

병원 간 네트워크에 대한 시각화의 경우 연구자들의 목적에 따라 다양한 형태가 필요하며, 선행 연구들 역시 목적에 따라 시각화를 도입하였다. 예를 들어 일부 선행 연구들은 네트워크 구조에 대한 시각화보다는 특정 질병에 대해 치료 경과가 어떻게 변화했는지에 대해 정량적으로 분석했다 [1], [5]. 해당 연구들은 특정 지역에 한정하여 진행되었으나, 다른 선행 연구들의 경우 분석의 범위를 국가 전체로 넓혀서 분석했다 [8], [9]. 아울러 지리적인 정보 외에도 네트워크 내에서 긴밀하게 서로 연결된 클러스터를 도출하여 분석을 한 사례도 있었다 [10].

본 연구에서는 병원 간의 네트워크에 대하여 시각적인 분석을 했던 사례들을 살펴보고, 뇌졸중 환자의 이동 데이터를 바탕으로 분석의 내용과 범위에 따라 효율적인 시각화 기법들을 제안한다. 특히 지리적인 정보를 활용하는 경우와 이를 배제하고 시각화하는 경우에 대해서 살펴봄으로써 각 방법의 장단점을 제시한다.

II. 관련 연구

환자의 이동에 대한 분석은 다양한 선행연구에서 진행되었다. 미국의 환자 이동 네트워크에서 병원군이 퍼지는 과정을 분석한 사례[8]가 있었으며, 지도 위에 2년간의 환자의 이동을 바탕으로 구성된 네트워크 구조를 노드-링크 다이어그램

형태로 표시하고, 링크의 색깔로 이동한 환자의 수를 표현했다. 그리고 이를 바탕으로 병원군의 전파를 감지할 수 있는 병원들을 네트워크 지표(i.e., in-degree)를 바탕으로 선택했다. Nekkabet al.[9]은 병원에서의 감염(HAI, hospital-acquired infections) 여부에 따라 환자 그룹을 3가지(with HAI, with suspected HAI, all patients)로 나누고, 해당 환자들의 이동 패턴을 바탕으로 네트워크 연결 구조를 도출하여 어떤 네트워크가 HAIs를 연구하기 위해 최적인지 살펴보았다. 특히 프랑스의 각 지역과 광역 행정 구역에서의 이송 패턴을 확인했으며 주요 병원을 찾을 때 중요한 역할을 할 수 있는 환자 이송 패턴을 확인했다. 해당 연구에서 네트워크는 지도상에 노드-링크 다이어그램 형태로 시각화했으며, Clauset[11]의 커뮤니티 검출 알고리즘을 사용하여 도출한 커뮤니티에 따라 노드에 색깔을 칠하여 커뮤니티를 구분했다. Dong et al.[10]의 연구에서도 경우 중국 내의 병원 간 네트워크를 분석하면서 커뮤니티 검출 알고리즘[12]을 사용했다. 다만 네트워크 분석 과정에서 지리적인 정보는 활용하지 않았으며, Gephi [13]를 사용하여 노드-링크 다이어그램 형태로 네트워크를 시각화 했다.

뇌졸중 환자의 치료와 관련하여 환자의 이동을 기반으로 한 연구도 진행되었다. Adeoye et al.[2]은 미국에서 뇌졸중 환자들의 지리적 접근성에 대해서 연구를 진행했는데, 국가 단위의 데이터 분석을 통해 치료를 받을 수 있는 병원에 대한 접근성과 실제 치료 비율을 비교함으로써, 보다 효율적인 뇌졸중 치료 시스템이 필요함을 주장했다. 이 때 치료 시설에 대한 접근성은 choropleth map을 사용하여 시각화 했다. Zachrisson et al.[7]은 네트워크에 대한 연구가 뇌졸중 치료 시스템을 개선하는데 활용될 수 있음을 주장했으며, 다른 의학 연구에서 커뮤니티 검출 알고리즘을 사용했던 사례를 소개했다. 그리고 해당 연구 이후 다양한 지역을 대상으로 네트워크 분석을 접목한 연구가 진행되었다. Zachrisson et al.[5]은 미국의 북동부 지역에서의 뇌졸중 환자 이동 데이터에 대해 분석했고, ArcGIS[14]를 사용하여 네트워크를 시각화 했다. 해당 연구에서는 두 시점(2007년과 2011년)의 환자 이동 데이터를 바탕으로 구성된 네트워크를 지도 상에 노드-링크 다이어그램 형태로 표현했으며, 환자를 많이 보내는 병원과 받는 병원에 대한 분석을 진행했다. 뇌졸중 치료 시스템에 대한 네트워크 분석은 California에 대해서도 진행되었다 [1]. EVT를 도입에 따른 변화를 확인하기 위하여 2013년부터 2014년 사이의 환자 이동을 기반으로 구성된 네트워크와, 동일한 방법으로 2016년부터 2017년 사이의 데이터를 기반으로 구성된 네트워크에 대해서 지도 위에 시각화를 했으며, EVT 치료를 진행 비율과 환자 이동 비율 등을 분석했다.

네트워크를 시각화 하는 방법은 자료의 크기나 시각화 목적에 따라 다양한 방법이 연구되었다. Komarek et al.[15]은 화면 상에 노드와 링크를 배치하는 layout 방식에 따라서 7가지(i.e., force-directed, hive plot, adjacency matrix, arc diagram, Sankey diagram, chord diagram, pivot

graph)로 분류를 했으며, 시각화하기에 적합한 노드와 링크의 수를 비교했다. Vehlow et al.[16]은 네트워크 내의 하위 그룹 구조를 시각화 한 선행 연구들에 대하여 문헌 조사를 진행하여 분석했다. 본 연구에서는 상술된 선행 연구에서의 시각화 기법들에 대해서 지리적 정보를 더했을 때 고려해야 하는 부분들과 커뮤니티 검출 결과를 나타낼 수 있는 방안들에 대하여 살펴본다. 특히 이 과정에서 한국 내에서 환자들의 이동 데이터가 가지고 있는 특성들을 반영한 효율적인 시각화 방법을 제시하고자 한다.

III. 연구 배경

3-1 데이터 소개

본 논문에서는 환자들의 이동 데이터를 기반으로 구성된 네트워크에 대한 시각화를 목표로 한다. 공동 연구를 진행 중인 연구 기관에서는 이를 위해 한국의 건강보험 심사평가원에서 제공하는 환자들의 이동 데이터를 활용하고 있다. 제공된 데이터는 특정 질병에 대한 치료 기관으로 등록된 병원들의 목록과 해당 병원 사이에서의 환자들의 이동, 그리고 환자들의 치료 과정과 그 결과에 관련된 데이터등이 있다. 이 중에서 공동 연구를 진행한 의학 연구자들이 다룬 데이터는 2016년 7월 1일부터 12월 31일 사이의 기간 동안 19,113명의 뇌졸중 환자들이 1009개의 병원 간에 이동한 정보로 구성되어 있다.

<표 1>에서는 해당 데이터 중에서 본 논문에서 활용한 데이터들을 정리했다. 병원의 고유 번호는 관리를 위해 부여한 번호들이며, 각 병원의 규모에 따라 상급종합병원, 종합병원, 병원, 의원으로 분류하고 있다. 병원의 주소는 건강보험 심사평가원에서 시군구 정보까지만 제공했기 때문에 병원을 특정하기 어려운 구조로 되어있으며, 지도상에 위치를 표시하기 위하여 해당 지역의 관청 소재지의 좌표를 병원의 위치로 같 음했다. 다음으로 SCH는 뇌졸중 집중 치료 병원(Stroke Care Hospital)의 지정 여부로써, 병원의 규모에 따른 분류 외에 치료를 위한 일정 여건이 갖춰져 있는지를 나타낸다.

제공되는 데이터들은 병원의 주소에서 확인할 수 있듯이 병원을 특정할 수 없도록 익명화되었다. 이 때문에 병원의 위치를 지도에 개략적으로 나타내는 과정에서 동일 행정 구역 내에 여러 개의 병원이 존재하는 경우 병원들이 겹쳐서 나타날 수 있는 문제가 발생했다. 이를 보완하기 위하여 지도에 렌더링하는 과정에서 픽셀 단위로 랜덤 지터링(jittering)을 적용하여 겹치지 않고 화면에 표시하도록 했다. 다만 병원 사이의 이동 거리를 산출할 때는 지터링이 적용되기 이전의 GPS 좌표들을 기준으로 하여 직선 거리로 환자들의 이동 거리를 산출했다.

환자의 이송 데이터 역시 환자 개인을 식별할 수 없도록 병원 간에 몇 명의 환자가 이동했는지 합계만 제공되었다.

표 1. 병원 및 환자 이동 데이터

Table 1. Hospital and Patient Transfer Data

Data	Item	Description	Note
Hospital Data	ID	Unique ID	—
	Type	Hospital types by size	4 types
	Address	Address	Restricted to City, county, and district information
	Location	GPS coordinates	Calculated from the government office location
	SCH	Designated SCH or Not	—
Patient Transfer Data	Type	Transfer type	“Stay” or “Transfer”
	Origin	Transfer origin hospital ID	—
	Destination	Destination hospital ID	—
	Patient Count	Number of the transferred patients	—

앞서 병원의 목록에 있었던 병원의 고유 번호를 바탕으로 출발지 병원과 도착지 병원, 그리고 이동한 환자의 수만 제공되었다. 환자의 이동 형태 중에서는 처음 방문한 병원에서 치료를 받고 그대로 퇴원한 경우(stay)와 처음 병원에서 치료를 받지 않고 다른 병원으로 한 번 옮겨서 치료를 받고 퇴원한 경우 (transfer) 두 가지를 다루었다. 실제 이동 형태는 이 밖에도 첫 병원에서 치료를 받고 다른 병원으로 이동한 후 퇴원한 경우와 첫 병원에서 두 번째 병원으로 이동 후 치료를 받고, 세 번째 병원으로 이동한 후 퇴원한 경우도 있었다. 다만 이러한 경우들은 환자 수를 모두 합쳐서 약 0.43%의 비율이었기 때문에, 본 연구에서는 가장 많은 비중을 차지하는 두 가지 형태의 이동만을 시각화하는 것을 목표로 했다.

3-2 목업(Mockup) 데이터 생성

위에서 언급한 것처럼 익명화를 위하여 여러 가지 과정을 거쳤음에도, 이번 연구에서 다루는 데이터는 민감 정보로 분류되어 연구 목적을 위하여 지정된 기관 외부로의 반출이 어려웠다. 이에 따라 본 연구의 경우 의학 연구자가 실제 데이터셋을 살펴볼 수 있는 환경과는 별도로 개발 및 검증에 사용할 수 있는 목업(mockup) 데이터를 준비하게 되었다. 목업 데이터의 경우 먼저 실제 데이터의 통계적 대략값을 바탕으로 행정 구역 당 병원의 수가 실제와 유사하도록 생성했고, 병원 간의 연결에서 환자를 보내는 병원과 받는 병원 역시 각 지역에 소재한 병원의 수를 가지고 산출한 지역별 병원 분포를 바탕으로 랜덤하게 생성했다. 또한, 해당 연결에서의 이동한 환자 수도 원본 데이터의 병원 간 이동 사례로부터 몇 명의 환자가 이동했는지에 대한 도수 분포를 조사하여, 이동한 환자 수의 출현 비율에 따라 랜덤하게 선택되도록 했다 (e.g., 10명의 환자가 이동한 사례가 전체 이동 사례의 10%일 경우, 10%의 확률로 10명의 환자가 이동한 것으로 설정).

3-3 시각적 분석 대상 분류

문헌 조사 및 의학 연구자들과의 면담을 통해 선행 연구에서 진행되었던 네트워크 구조 분석 과정에서의 시각화 대상들을 분류해보았다. 먼저 시각화할 데이터의 지정학적인 규모를 대상으로 살펴보자면, (1) 국가 단위의 네트워크를 구성하고 시각화 하는 것을 상정할 수 있다. 이 경우 전체 데이터에 대해서 별다른 필터링 없이 모두 시각적 분석의 대상으로 삼게 된다. 다음으로는 (2) 하위 행정 구역 단위로 시각화 대상을 한정하는 경우이다. 이렇게 분석 대상이 전체 데이터 중 일부분으로 한정되는 경우, 해당 지역 외에 다른 지역으로의 환자 이동의 처리 방식에 따라 행정 구역 내에서의 이동만 고려하거나, 해당 행정 구역과 환자를 주고받은 다른 행정 구역들도 함께 시각화하는 방안이 있다. 전자의 경우 출발지와 도착지가 특정 지역 내에 모두 존재하는 경우만 살펴보기 때문에 시각적인 어수선했음(visual clutter)을 상대적으로 줄일 수 있으며, 지역 내의 네트워크 구조에 집중할 수 있다. 후자의 경우에는 지역 간의 이동에 대해서도 시각화를 하기 때문에, 지역 간의 불균형 등 주목할만한 경향성이 있는지 살펴볼 수 있다. 다음으로는 행정 구역에 따른 데이터의 부분 집합을 시각화하는 것이 아니라, (3) 네트워크 내에 존재하는 노드들을 군집화하고, 이 결과를 바탕으로 군집 단위로 시각화하는 방안이 있다. 이 경우 노드들 사이의 연결 관계에 따라 행정 구역을 넘나드는 군집이 도출될 수도 있게 된다.

네트워크 시각화 과정에서는 노드 외에도 노드 사이의 연결에 대한 시각화 역시 분석의 목적에 따라 결정되어야 한다. 본 연구에서 목표로 하는 환자들의 이동에 대한 시각화의 경우 이동의 방향성이 존재한다. 따라서 (1) 링크를 표시하지 않는 경우, (2) 링크를 표시하되 방향성은 생략하고 연결 여부만 표시하는 경우, (3) 연결의 방향성을 함께 표시하는 경우가 존재한다. 추가로 본 연구에서 다루고 있는 데이터에는 환자가 이동하지 않고 한 병원에서 머문 경우가 있었기 때문에, 이러한 셀프-루프(self-loop)를 시각화의 범위에 포함할지 여부 역시 선택이 필요하다. 아울러 각 연결의 경우 이동한 환자의 숫자가 존재하기 때문에, 모든 정보를 포함할 경우, 방향성이 있는 가중치 그래프의 형태가 된다.

3-4 사용자 과업 분석

본 연구는 뇌졸중 환자들에 대해서 병원 간 네트워크 구성을 연구하는 의학 연구자들을 위한 시각화를 제안하는 것을 목표로 했다. 해당 사용자들의 분석을 지원하기 위하여 선행 연구에 대한 문헌 조사와 함께 대상 연구자들에 대한 인터뷰를 진행했다. 참여한 연구자들은 신경과 전문의 2명과 임상 통계학자 1명으로, 다음의 대표적인 분석 과업을 확인할 수 있었다.

연구자들이 언급한 주요 과업 중 하나는 군집 단위의 네트워크 구조 분석이었다. EVT의 도입으로 인하여 병원의 군집화와 관련된 연구가 의학계에서 진행되기 시작했다. 이에 따

라 군집을 확인할 수 있는 정량적인 방법들과 이를 시각적으로 확인할 수 있는 방법들이 연구되었다. 먼저 정량적인 분석의 경우 네트워크의 연결 관계를 확인할 수 있는 정량적 지표(e.g., centrality 또는 density)를 분석하는 방법이 제시되었으며, 시각적으로 이를 확인하기 위하여 노드-링크 다이어그램 기반의 시각화 기법들이 활용되기도 했다. 인터뷰에 참여한 연구자들은 두 가지 형태의 분석을 모두 주요 과업으로 상정했으나, 본 논문에서는 이 중에서 시각적 분석을 지원하는 것을 목표로 했다.

다음으로 연구자들이 꼽은 대표적인 과업은 지역 단위의 환자 이동과 지역 간의 관계 분석이었다. 한국에서는 질병의 치료를 위하여 진료권의 설정과 권역 내의 거점 추가 등 효율적인 진료권 운영 방안에 대해서 연구를 진행하고 있지만, 진료권의 설정 기준이나 방법에 있어서도 다양한 대안들이 제시되고 있다. 협업을 진행한 연구자들은 이렇게 설정된 권역이 잘 동작하고 있는지에 대한 간접적인 확인을 위해서 지역 단위의 환자 이동 패턴을 분석하고자 했다. 예를 들어 한국의 일부 지역에서는 특정 병원에 대한 의존도가 높고, 인근 지역으로 환자를 보내는 사례도 많다고 통상적으로 알려져 있는데, 이런 부분에 대해서 실제로 살펴볼 수 있으면 장기적으로 진료권을 설정하고 개선하는데 있어서 도움이 될 것으로 기대했다. 실제로 선행 연구 중에서는 이를 위해 병원의 위치 정보를 바탕으로 지도상에 병원들을 배치하고 연결 관계를 나타낸 사례들이 있었으며, 앞서 언급된 군집 단위의 분석에서 나아가 지역 단위의 환자 이동과 지역 간의 관계에 대한 분석도 수행하는 것을 목표로 했다.

IV. 병원 간 네트워크 시각화

본 논문에서는 다양한 병원 간 네트워크의 시각화를 웹 기반으로 구현했다. Komarek et al.[15]이 분류한 7가지(i.e., force-directed, hive plot, adjacency matrix, arc diagram, Sankey diagram, chord diagram, pivot graph)의 네트워크 시각화 방법 중에서 본 논문에서는 force-directed, chord diagram 및 adjacency matrix를 기반으로 한 시각화들을 제안하며, 이 외의 시각화 방법에 대해서는 향후 연구에서의 활용 방안에 대하여 다룬다. 시각화에는 D3.js[17]를 사용했으며, 지도 상에서 노드들의 군집을 표현하기 위하여 Bubble Sets[18]의 JavaScript 라이브러리를 사용했다. 시각화 결과물은 HTML5와 CSS3 기반의 페이지에 SVG 1.1로 출력되었다.

4-1 Node-link 다이어그램을 활용한 시각화

사용자의 주요 과업 중 하나로 확인된 군집 단위의 네트워크 구조 분석을 위하여 네트워크의 연결 관계 네트워크 시각화 방법 중 노드-링크 다이어그램을 활용하여 시각화를 진행

했다. 여기에서는 별도의 군집화 알고리즘은 적용하지 않았으나, force-directed 레이아웃 알고리즘을 통해 노드들을 배치함으로써 긴밀하게 연결된 노드들이 모이는 것을 확인할 수 있었다(그림 1). 노드-링크 다이어그램의 경우 링크의 특성을 바탕으로 방향성과 가중치를 표시할 수 있으나, 1000개 내외의 노드들과 2000개 내외의 링크로 구성된 네트워크 규모에서 방향성을 표시하는 것은 시각적 혼란을 일으킬 수 있었기 때문에 방향성 없이 시각화했다. 이에 따라 한 방향으로의 환자 이동만 있더라도 노드 사이를 연결했으며, 두 병원 사이에서 이동한 전체 환자 수에 따라 두께를 다르게 설정했다. 아울러 군집을 시각적으로 확인하는 과정에서 군집이 각 병원이 속한 지역에 따라 형성이 되어있는지 확인하기 위하여 각 노드의 색깔은 지역에 따라서 배정했다. 그 결과 군집이 형성되어 있음을 확인할 수 있었고, 대체로 같은 지역에 있는 병원들이 군집을 형성함은 확인할 수 있었지만, 지역 간의 관계나 다수의 행정 구역에 걸쳐있는 군집들을 확인하기에는 어려움이 있었다.

이러한 단점을 보완하기 위하여 위치 정보를 바로 확인할 수 있도록 지도 위에 노드를 배치한 것이 (그림 2)의 시각화이다. 먼저 위치 정보만을 활용해서 배치한 노드들 사이를 환자 이동 정보에 따라 연결해본 결과, 주요 도시에 병원들이 배치되어 있고, 해당 도시 인근의 병원들과 스타 토폴로지(star topology)를 구성하고 있음을 확인할 수 있었다. 본 연구에서는 실제로 군집의 분포를 확인하기 위하여 Louvain 알고리즘[19]을 통해 군집화를 수행했다.

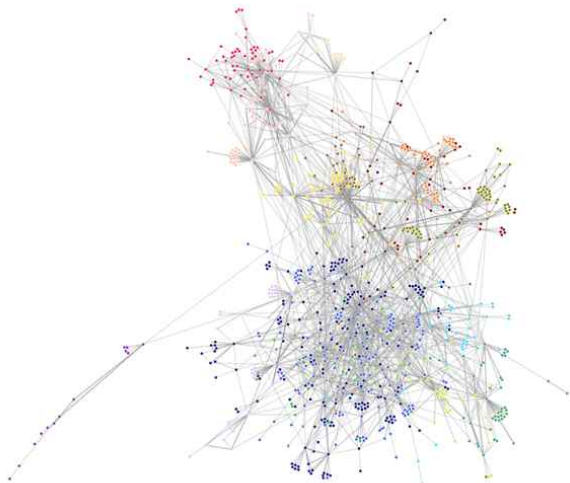


그림 1. Force-directed 레이아웃 알고리즘을 통해 노드-링크 다이어그램으로 시각화한 결과. 각 노드의 배경 색깔은 병원이 속한 지역에 따라 결정되어 있으며, 이동한 환자의 숫자에 따라 링크의 두께를 설정

Fig. 1. Node-link visualization with force-directed layout algorithm. Each node is colored by its province, and the width of the link corresponds to the number of transferred patients

도출된 군집들은 (그림 2b)와 같이 Bubble Sets[18] 형태로 표현하여 지도 위에 나타내었다. 이는 도출된 군집의 수가 90개 남짓으로 색깔로 모두 구분하기에는 어려움이 있었기 때문에, 군집의 영역을 외곽선으로 둘러싸인 반투명한 영역으로 표현함으로써 유사한 색깔의 군집이 겹치는 지역이 있더라도 어느 정도는 구분할 수 있도록 했다. 그 결과 멀리 떨어진 행정 구역들에 속한 병원들로 구성된 이질적인 군집을 시각적으로 확인할 수 있었다.

4-2 Chord 다이어그램을 활용한 시각화

병원 간의 연결 관계를 노드-링크 다이어그램 형태로 표현하는 방법은 군집의 확인과 지역별 분포를 시각적으로 분석하는 데는 도움이 되었으나, 행정 구역 단위로 이동한 환자들의 수나 병원별로 거쳐 간 환자들의 수에 대한 정보는 확인하기 어려웠다. 이를 보완하기 위하여 Chord 다이어그램 기반의 시각화를 도입하게 되었다(그림 3).

원을 따라 배치된 호의 색깔은 16개의 광역 행정 구역에 따라 설정했으며, 두 개의 호의 쌍을 잇는 연결은 출발지의 행정 구역에서 이동하는 환자의 수에 따라 두께가 결정되었다. 따라서 호의 길이는 전체 환자들의 수에 대한 해당 행정 구역에서 발생한 환자의 수의 비로 결정되었다.

먼저 광역 행정 구역 단위로 환자들의 이동을 확인하기 위하여 병원 단위로 chord diagram을 통한 시각화를 적용한 결과, (그림 3a)와 같은 결과를 얻을 수 있었다.

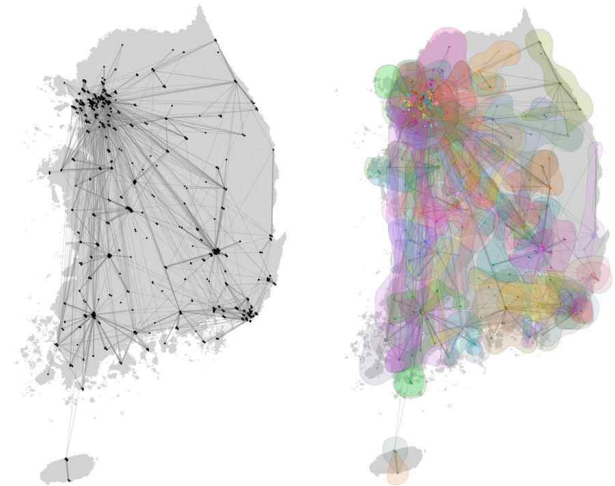


그림 2. 각 노드를 주소에 따라 지도에 배치한 시각화 결과 (좌측) 별도의 시각적 인코딩 없이 병원의 위치와 연결 관계만을 시각화 (우측) 군집화를 했을 때 도출되는 군집에 따라 색깔을 배정하고, Bubble Set으로 표현한 시각화 결과

Fig. 2. Visualization of nodes and links on a map. (Left) Nodes are in black dots and gray lines corresponds to patient transfers. (Right) Visualization of groups from community detection algorithm, each colored and presented using Bubble Sets.

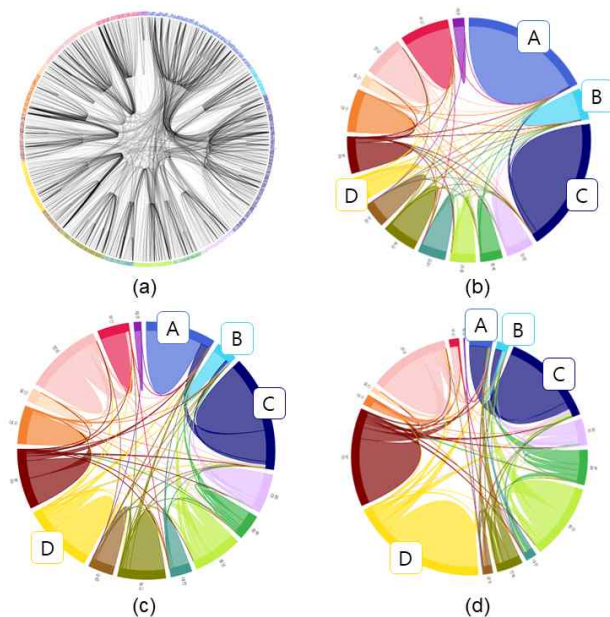


그림 3. Chord diagram 기반의 네트워크 시각화 결과. (a) 병원 단위로 나타난 다른 병원으로 이동한 환자들에 대한 데이터 (b) 지역 단위로 묶은 전체 환자 데이터 (c) 다른 병원으로 이동한 환자들에 대한 데이터 (d) 다른 행정 구역으로 이동한 환자들에 대한 데이터.

Fig. 3. Chord diagram visualizations of network with (a) patients transferred to other hospitals for each hospital, (b) whole patient data grouped by province, (c) patients transferred to other hospitals grouped by province, and (d) patients transferred to other provinces.

광역 행정 구역에 속한 병원들을 모아서, 같은 색깔로 병원의 고유 ID를 호 외부에 배치했고, 시각적 혼란을 줄이기 위하여 비슷한 방향으로의 이동을 나타내는 링크들을 모아주는 계층적 엣지 번들링(hierarchical edge bundling)을 적용했다. 이를 통해 인접한 행정 구역 간의 연결과 단일 행정 구역 내의 이동을 확인할 수 있었지만, 전체 노드의 수가 1000개가 넘었기 때문에 행정 구역 간의 관계를 확인하는 데 어려움이 있었다.

이에 따라 이러한 한계를 극복하기 위하여 광역 행정 구역 단위로 환자들을 집계한 시각화를 도입하게 되었다. 한 병원에서 머문 환자들을 포함한 전체 환자 데이터를 시각화한 결과가 (그림 3b)로, 지역별로 환자의 발생 비율을 효과적으로 확인할 수 있었다. 그림의 오른쪽 상단에 있는 서울(A), 인천(B), 경기(C) 지역 등 수도권에서 발생한 환자의 비율이 월등히 높음을 확인할 수 있다. 하지만 이러한 시각화는 지역 내에 머문 환자들의 수가 많은 경우 상대적으로 다른 지역으로의 이동을 확인하기 어렵다는 단점을 확인할 수 있었다. 이처럼 한 지역에 환자가 머무는 상황은 다음의 두 가지로 (1) 한 병원에서 머물다가 퇴원하거나 (2) 동일 행정 구역 내에 존재하는 다른 병원으로 이동한 후 퇴원하는 상황이다. 이러한 환자들은 Chord 다이어그램 상에서 자신으로 향하는 링크(셀프-루프, self-loop)가 되어 둥근 부채꼴 형태로 표시된다.

본 연구에서 사용한 데이터는 이러한 두 종류의 셀프-루프 중에서도 특히 한 병원에 머물다 퇴원한 환자의 수가 상대적으로 많았기 때문에, 둥근 부채꼴 형태로 표시되는 구역의 크기가 매우 커져서 (그림 3b)와 같이 행정 구역간의 환자 이동 비율을 확인하기 어려웠다.

상술한 바와 같이 한 병원에 머물다 퇴원한 환자로 인해 다른 환자들의 비율을 확인하기 어려운 단점을 극복하기 위하여, 전체 데이터 중에서 다른 병원으로 이동한 환자들만 나타낸 결과 (그림 3c)와 같은 패턴을 확인할 수 있었다. 서울(A)의 경우 다른 지역과 비교하여 환자의 비율이 많이 줄어든 것을 확인할 수 있었는데, 이를 통해 서울의 경우 한 병원에서 치료를 받고 퇴원한 환자들의 비율이 매우 높음을 알 수 있다. 또한, 다른 병원의 이동만 고려했을 때도 셀프-루프의 비중이 높음이 확인되었는데, 이는 서울(A)이나 경기(B)의 경우, 다른 지역으로 이동하지 않고 해당 행정 구역 내에 위치한 다른 병원으로 이동한 환자의 비율이 높음을 의미한다. 다음으로 서울(A)과 경기(C) 사이에서는 비대칭적인 환자의 이동을 확인할 수 있었다. 두 지역을 잇는 리본의 경우 서울(A)쪽은 두께가 얇고, 경기(C) 지역의 두께가 매우 두꺼웠는데, 이를 통해 서울에서 경기로 이동하는 환자의 수가 반대 방향으로 이동하는 환자의 수보다 상대적으로 적음을 알 수 있다.

마지막으로 (그림 3d)의 경우 한 지역 내에서만 있었던 환자들(i.e., 한 병원에서 치료를 받고 퇴원했거나, 같은 지역 내의 다른 병원으로 이동한 후 퇴원한 환자들)을 제외하고, 다른 행정 구역으로 이동한 환자들을 시각화한 결과이다. 여기에서도 서울(A)과 경기(C) 사이의 비대칭적 환자 이동 패턴을 확인할 수 있었다. 아울러 전남(D)지역의 환자 비율이 이전에 살펴본 시각화 결과와 비교했을 때 매우 높아진 것을 확인할 수 있었으며, 매우 많은 수의 환자가 인접 행정 구역으로 이동했음을 알 수 있다.

4-3 Adjacency matrix 기반의 시각화

앞서 살펴본 Chord diagram 기반의 시각화는 광역 행정 구역 단위의 관계를 살펴보기에는 적합했으나, 데이터를 병합하여 시각화한 관계로 병원 단위의 분석을 진행하기에는 어려움이 있었다. 사용자 과업 분석 과정에서 확인된 분석 요구 사항 중 하나는 군집에 대한 시각적 분석과 권역 내의 거점의 확인이었는데, 본 연구에서는 이러한 분석을 지원하기 위하여 adjacency matrix를 기반으로 한 시각화를 도입했다.

전체 환자 이동 데이터를 시각화한 결과 (그림 4)와 같은 패턴을 확인할 수 있었는데, 행과 열은 병원에 대응되며, 행이 출발지 병원, 열이 도착지 병원을 나타낸다. 따라서 대각선에 있는 셀들의 경우 다른 병원으로 이동하지 않고 해당 병원에서 치료받은 환자들을 표시하게 된다. 각 셀에는 해당 병원들 사이에 이동한 환자의 수에 따라 환자의 수가 많을수록 농도를 진하게 색깔을 칠했으며, 도착지 병원의 행정 구역에 따라 색깔을 지정했다.



그림 4. Adjacency matrix 기반의 병원 간 환자 이동 시각화. 행과 열은 병원에 대응되며, 셀의 색깔은 도착 병원의 행정 구역에 따라 결정된다. 색의 농도는 이동한 환자의 수에 따라 많을수록 진하게 표시했다.

Fig. 4. Visualization of patient transfers between hospitals based on adjacency matrix. Each row and column corresponds to each hospital, and the color of each cell is determined from the province of the destination hospital. The saturation of each cell corresponds to the number of transferred patients.

그 결과 세로 방향으로 색깔이 진한 점이 찍힌 병원들을 확인할 수 있었는데, 이는 여러 병원으로부터 하나의 병원에 환자들이 이동했음을 의미한다. 위의 결과에서는 크게 두 가지 형태의 세로로 긴 패턴이 있었는데, 첫 번째는 (그림 4A)와 같이 한 지역 내에서 특정 병원으로의 이동이 많은 경우이다. 이와는 달리 (그림 4B)에 있는 병원들의 경우 같은 지역으로부터의 환자 이동뿐 아니라 다른 지역으로부터의 환자 이동이 많았기 때문에 다른 색깔의 점들로 구성된 세로로 긴 패턴이 확인되었다. 두 경우 모두 거점의 역할을 하는 병원임을 추론할 수 있으나, 전자의 경우 행정 구역 내에서 군집이 도출될 확률이 높지만, 후자의 경우 두 개 이상의 행정 구역에 걸쳐 군집이 도출될 수 있기에, 권역에 대한 추가적인 분석이 필요함을 확인할 수 있다. 또한, 인접하지 않은 행정 구역으로부터의 환자 이동(그림 4C)도 확인할 수 있었는데, 이러한 이동의 경우 권역을 벗어난 상대적으로 먼 거리의 이동을 의미하기 때문에, 빠른 치료가 우선시되는 뇌졸중의 특성상 특이한 이동 패턴임을 알 수 있다. 아울러 본 연구에서 다른 데이터의 경우, 병원의 수가 1000개를 넘는 관계로 각 행과 열의 수가 많아서, 상술한 방법으로는 개괄적인 패턴만 확인할 수 있었다.

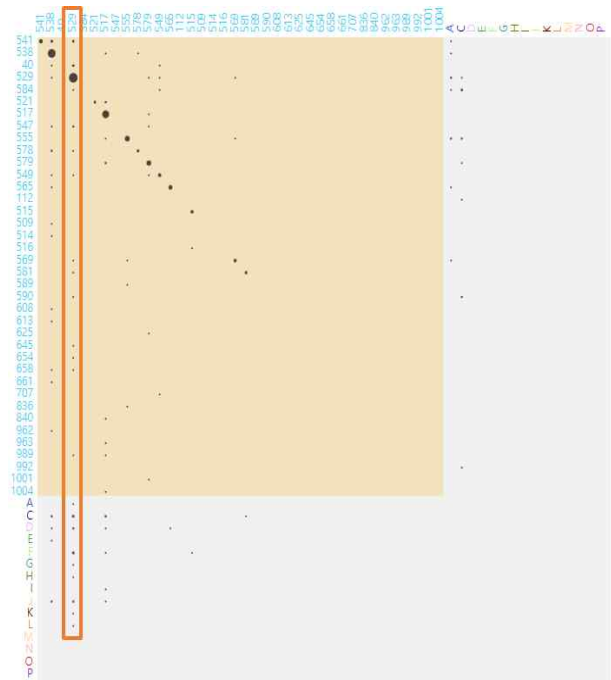


그림 5. 특정 행정 구역에 대한 adjacency matrix 기반의 병원간 환자 이동 시각화 결과. 옅은 노란색 배경은 행정 구역 내의 병원들을 나타내는 영역을 표시하며, 숫자는 병원의 ID를 나타내고, 알파벳은 다른 행정 구역을 나타낸다.

Fig. 5. Visualization of patient transfers for a specific province based on adjacency matrix. Pale yellow background depicts hospitals within the province, the numbers on rows and columns correspond to hospital, and alphabets on rows and columns denotes other provinces.

이에 따라 세부적인 패턴을 확인할 수 있도록 마우스 휠 스크롤을 통한 동적인 확대 및 축소를 구현했으며, 셀에 마우스 커서를 올려놓을 경우, 관련된 병원들의 ID와 이동한 환자의 수가 화면에 나타나도록 함으로써, (그림 4의 A, B, C)와 같은 패턴을 나타내는 병원들의 ID와 환자의 수에 대한 정보를 확인할 수 있었다.

다음으로는 앞서 살펴본 전국 단위의 시각화로부터 더 나아가 지역 단위의 분석을 위한 (그림 5)와 같은 시각화를 구성했다. 여기서는 각 행과 열에 표시된 숫자들은 병원의 ID를 나타내고, 알파벳은 해당 지역을 제외한 나머지 행정 구역에 대응되도록 구성했다. 또한, 행정 구역 내에 속한 병원들인지를 명확하게 확인할 수 있도록 옅은 노란색 배경으로 행정 구역 내의 병원들을 구분해주었다. 앞서 살펴본 (그림 4)의 시각화와 동일하게 행이 출발지, 열이 도착지를 나타내도록 구성했는데, (그림 4)의 시각화와는 달리 각 셀에는 색깔의 농도가 아닌 원의 크기로 환자의 수를 표시하도록 했다. 이는 행정 구역 내를 표시하는 배경 색깔에 의한 혼동을 막기 위한 것으로, 정확한 환자 숫자의 확인을 위해서 마우스 커서를 셀 위에 올려두는 상호 작용을 구현함으로써 실제로 이동한 환

자의 수를 확인할 수 있도록 했다. 그 결과 (그림 5)에서 주황색 테두리로 표시된 병원이 같은 행정 구역 내로부터 환자를 많이 받을 뿐만 아니라 다른 행정 구역으로부터 들어오는 환자들이 많이 있음을 확인할 수 있었다.

V. 논의 및 결론

다양한 선행 연구들이 네트워크 구조를 표현하는 방법을 연구했으나, 본 연구에서는 Komarek et al.[15]이 분류한 7가지 중 hive plot, arc diagram, Sankey diagram, pivot graph를 제외한 나머지 시각화 방안들을 활용했다. 본 장에서는 먼저 논문에서 다룬 시각화들을 제외한 나머지 시각화 방법들에 대해 논의하고자 한다. Hive plot의 경우 방사형으로 배치된 여러 개의 축에 노드들을 배치하고, 노드 사이의 연결을 호를 통해 표시한다. 각 축은 네트워크들을 구성하는 노드들에 대응되는 정량적인 자료가 있을 때, 해당 자료를 기준으로 결정하게 되는데, 표시할 축의 수와 길이를 조절함으로써 많은 수의 노드들에 대해서도 효과적인 시각화가 가능하다. 하지만 본 연구에서 다룬 데이터의 경우 노드들에 대한 다변량 데이터가 아니었기 때문에, 해당 방법을 적용하지 않게 되었다. 다음으로 arc diagram과 Sankey diagram의 경우 선행 연구[15]에 따르면 수백 개 정도의 노드들이 있을 때 시각적 혼란 없이 표현할 수 있게 되는데, 이는 본 연구의 규모와 비교했을 때 부족함이 있어 선택하지 않게 되었다. 마지막으로 Pivot graph의 경우 각 노드가 여러 가지 특성을 갖는 다변량 데이터에 적합하고, 효율적으로 표현할 수 있는 노드의 수가 수십 개에 불과했기 때문에 활용하지 않았다.

한편, 노드의 수와 관련된 한계를 극복하기 위해 데이터를 가공하는 방안이 있는데, 본 연구에서는 병원들이 속한 지역별로 환자 데이터를 병합하여 지역 간의 관계를 분석하는데 이를 활용했다. 이를 통해 실제로 몇 개의 지역에 대해서 차별화된 특징을 발견할 수 있었으며, 의료 권역 설정 및 운영 관점에서의 연구에 대한 가능성을 확인할 수 있었다. 향후 연구에서는 행정 구역 외에도 다양한 군집화 알고리즘을 적용한 결과를 바탕으로, 병원들을 대신하여 군집들 사이의 관계를 확인해봄으로써 의학 연구로 이어질 기회를 마련할 수 있을 것으로 기대한다. 또한, 이 경우 각 병원은 행정 구역 외에 군집에 대한 포함 관계를 갖게 되기 때문에 다변량 데이터에 특화된 다른 시각화 기법을 활용할 수 있을 것으로 기대된다.

본 연구에서는 환자들의 병원 간 이동 데이터를 기반으로 네트워크를 구성하고, 해당 네트워크를 시각화하는 다양한 방법들을 살펴봤다. 이 과정에서 사용자 과업 분석을 통해 의학 연구자들이 관심을 갖고 있는 분석에 대한 수요를 확인했으며, 선행 연구에서 제시된 시각화 기법들을 기반으로, 이러한 요구를 충족할 수 있는 데이터 전처리 방법 및 시각화 방법에 대해 다루었다. 이 과정에서 전체 분석 과정을 모두 지원할 수 있는 시각화를 구성하기보다는 여러 형태의 시각화를 연

계함으로써, 데이터를 여러 각도로 분석하는 방안을 채택했다. 향후 연구에서는 각 병원 및 행정 구역 중심의 시각적 분석에서 나아가 군집 단위의 분석을 효율적으로 활용할 수 있는 방안을 도출하고, 네트워크나 하위 군집들의 구조에 대한 정량적인 지표들에 대해서도 분석을 진행해볼 수 있을 것으로 기대한다. 아울러 병원이나 환자들에 대한 추가적인 의료 정보를 확보함으로써 다변량 데이터를 구성하여, 본 연구에서 활용하지 못한 다변량 데이터에 대한 다양한 시각화 기법들의 도입도 가능할 것이다.

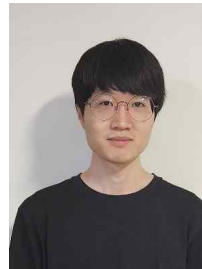
감사의 글

이 논문은 2021년도 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2021R1C1C1009327).

참고문헌

- [1] K. S. Zachrison, R. Y. Hsia, S. Li, M. J. Reeves, C. A. Camargo Jr, Z. Yan, J. P. Onnela, and L. H. Schwamm, "Ischemic Stroke Systems of Care in California: Evolution in the Organization During the Mechanical Thrombectomy Era," *Stroke: Vascular and Interventional Neurology*, p. e000206, July 2022. <https://doi.org/10.1161/SVIN.121.000206>
- [2] O. Adeoye, K. C. Albright, B. G. Carr, C. Wolff, M. T. Mullen, T. Abruzzo, A. Ringer, P. Khatri, C. Branas, and D. Kleindorfer, "Geographic access to acute stroke care in the United States," *Stroke*, vol. 45, no. 10, pp. 3019-3024, August 2014. <https://doi.org/10.1161/STROKEAHA.114.006293>
- [3] O. A. Berkhemer, P. S. Fransen, D. Beumer, L. A. Van Den Berg, H. F. Lingsma, A. J. Yoo, W. J. Schonewille, J. A. Vos, P. J. Nederkoorn, and M. J. Wermer, "A randomized trial of intraarterial treatment for acute ischemic stroke," *New England Journal of Medicine*, vol. 372, pp. 11-20, January 2015. <https://doi.org/10.1056/NEJMoa1411587>
- [4] M. Goyal, A. M. Demchuk, B. K. Menon, M. Eesa, J. L. Rempel, J. Thornton, D. Roy, T. G. Jovin, R. A. Willinsky, and B. L. Sapkota, "Randomized assessment of rapid endovascular treatment of ischemic stroke," *New England Journal of Medicine*, vol. 372, no. 11, pp. 1019-1030, March 2015. <https://doi.org/10.1056/NEJMoa1414905>
- [5] K. S. Zachrison, J.-P. Onnela, A. Hernandez, M. J. Reeves, C. A. Camargo Jr, M. Cox, R. A. Matsouaka, J. P. Metlay, J. N. Goldstein, and L. H. Schwamm, "Ischemic stroke transfer patterns in the northeast United States," *Journal of Stroke and Cerebrovascular Diseases*, vol. 28, no. 2, pp. 295-304,

- February 2019.
<https://doi.org/10.1016/j.jstrokecerebrovasdis.2018.09.048>
- [6] H. H. Lee, "Significance and limitations of the 2nd Basic Public Health Care Plan from the Perspective of Expanding Public Health Care," *Public Health Affairs*, vol. 5, no. 1, December 2021. <https://doi.org/10.29339/pha.21.4>
- [7] K. S. Zachrisson, A. Dhand, L. H. Schwamm, and J.-P. Onnela, "A network approach to stroke systems of care," *Circulation: Cardiovascular Quality and Outcomes*, vol. 12, no. 8, p. e005526, August 2019.
<https://doi.org/10.1161/CIRCOUTCOMES.119.005526>
- [8] J. Fernández-Gracia, J.-P. Onnela, M. L. Barnett, V. M. Eguíluz, and N. A. Christakis, "Spread of pathogens in the patient transfer network of US hospitals," in *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, Washington DC: DC, pp. 271-280, June 2017.
https://doi.org/10.1007/978-3-319-60240-0_33
- [9] N. Nekkab, P. Astagneau, L. Temime, and P. Crepey, "Spread of hospital-acquired infections: A comparison of healthcare networks," *PLoS computational biology*, vol. 13, no. 8, p. e1005666, August 2017.
<https://doi.org/10.1371/journal.pcbi.1005666>
- [10] X. Dong, D. Tang, and C. Tang, "Social network analysis in China's hospital healthcare," *Physica A: Statistical Mechanics and its Applications*, vol. 565, p. 125546, March 2021. <https://doi.org/10.1016/j.physa.2020.125546>
- [11] A. Clauset, "Finding local community structure in networks," *Physical review E*, vol. 72, no. 2, p. 026132, August 2005. <https://doi.org/10.1103/PhysRevE.72.026132>
- [12] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, p. P10008, October 2008.
<https://doi.org/10.1088/1742-5468/2008/10/P10008>
- [13] Gephi [Internet]. Available: <https://gephi.org/>.
- [14] ArcGIS [Internet]. Available: <https://www.arcgis.com/>.
- [15] A. Komarek, J. Pavlik, and V. Sobeslav, "Network visualization survey," in *Computational Collective Intelligence*, Springer, pp. 275-284, October 2015.
https://doi.org/10.1007/978-3-319-24306-1_27
- [16] C. Vehlou, F. Beck, and D. Weiskopf, "Visualizing group structures in graphs: A survey," in *Computer Graphics Forum*, vol. 36, no. 6, pp. 201-225, September 2017.
<https://doi.org/10.1111/cgf.12872>
- [17] M. Bostock, V. Ogievetsky, and J. Heer, "D³ data-driven documents," *IEEE transactions on visualization and computer graphics*, vol. 17, no. 12, pp. 2301-2309, December 2011. <https://doi.org/10.1109/TVCG.2011.185>
- [18] C. Collins, G. Penn, and S. Carpendale, "Bubble sets: Revealing set relations with isocontours over existing visualizations," *IEEE transactions on visualization and computer graphics*, vol. 15, no. 6, pp. 1009-1016, November/December 2009.
<https://doi.org/10.1109/TVCG.2009.122>
- [19] N. Dugué and A. Perez, "Direction matters in complex networks: A theoretical and applied study for greedy modularity optimization," *Physica A: Statistical Mechanics and its Applications*, vol. 603, p. 127798, October 2022. <https://doi.org/10.1016/j.physa.2022.127798>



곽규한(Kyuhan Kwak)

2021년 : 숭실대학교 대학원 입학

2021년~현재 : 숭실대학교 컴퓨터학부 석사과정

※ 관심분야 : 인간-컴퓨터 상호작용(Human-Computer Interaction), 상호작용 기법, 정보 시각화 등



박진우(Jinu Park)

2020년 : 숭실대학교 입학

2020년~현재 : 숭실대학교 컴퓨터학부 학사과정

※ 관심분야 : 인간-컴퓨터 상호작용(Human-Computer Interaction), 시각적 분석, 정보 시각화, 웹 서비스 등



송현주(Hyunjo Song)

2009년 : 서울대학교 컴퓨터공학부 (공학사)

2011년 : 서울대학교 전기·컴퓨터공학부 대학원 (공학석사)

2016년 : 서울대학교 전기·컴퓨터공학부 대학원 (공학박사-휴먼-컴퓨터 인터랙션)

2017년~2020년 : 덕성여자대학교 IT미디어공학과 조교수

2020년~현재 : 숭실대학교 컴퓨터학부 조교수

※ 관심분야 : 휴먼-컴퓨터 인터랙션(Human-Computer Interaction), 데이터 시각화, 시각적 분석 등