Healthcare Network Designing Using Graph Neural Networks

Ajay Kumar[1] Rajesh Kumar[2]

[1]kumar.327@iitj.ac.in
[2]kumar.332@iitj.ac.in

Abstract

The healthcare system is an intricate web of interrelated components, encompassing individuals seeking care, medical professionals, healthcare facilities, and insurance entities. This elaborate structure can be effectively represented using graph neural networks (GNNs). GNNs are a sophisticated subset of machine learning algorithms, exceptionally adept at handling data that encapsulates relationships. By employing GNNs, it's possible to decipher the intricate connections among the various elements within the healthcare framework. The insights gleaned from this analysis can then be harnessed to enhance both the caliber and the operational efficiency of healthcare services. We introduce an innovative method for leveraging graph neural networks (GNNs) to map out the intricate connections among patients, healthcare providers, hospitals, and insurers. The insights derived from this process inform the development of a revamped healthcare network, optimized for both efficiency and efficacy. The primary objectives of this new network are to curtail healthcare expenses and elevate the standard of patient care. To assess the viability of our methodology, we conduct an analysis on healthcare claims data. The findings indicate that our strategy has the potential to forge healthcare networks that surpass the efficiency and effectiveness of conventional models.

Keywords— Healthcare, Network Design, GNN, Healthcare Insurance

1 Introduction

1.1 Motivation

The healthcare sector is grappling with a multitude of issues, such as escalating expenses, growing intricacy, and a demographic shift towards an older population. These pressing concerns underscore the urgent need for inventive approaches that enhance the efficiency and quality of healthcare services. Graph neural networks (GNNs) stand out as a potent instrument for depicting the elaborate interplay among various healthcare entities, rendering them ideal for tackling the challenges in the healthcare domain.

Analyzing healthcare claims data can unveil recurring patterns and healthcare trends, which can be pivotal in shaping new healthcare strategies and policy reforms.

1.2 Healthcare Market Analysis

The global healthcare market is a massive and ever-growing industry, worth an estimated **$8.4 trillion in 2023** and projected to reach **$11.9 trillion by 2028**. This growth is being driven by several factors, including the aging population, rising chronic diseases, and increasing demand for personalized medicine. In addition, the **private health insurance industry**significantly influences the healthcare market by providing coverage to millions of individuals worldwide. The operational expenses incurred by these insurers are impacted by various factors, including the design of health plans, the health history, age, and overall well-being of enrolees. These operational costs are subsequently passed on to consumers in the form of **high premiums**. The high cost of healthcare is a major concern for many people around the world. In the United States (*which is a more mature IT healthcare market*), for example, the average annual premium for health insurance is over $10,000 per person [**?**]. This can make it difficult for people to afford the care they need. There are several factors that are contributing to the high cost of healthcare, including the rising cost of drugs and technology, the increasing complexity of medical care, and the growing demand for services.

1.3 Healthcare Terminologies

Payor: A health care provider encompasses both individual practitioners and healthcare facilities authorized to offer diagnostic and treatment services. This includes professionals licensed to prescribe medications, perform surgical procedures, and utilize medical devices, all aimed at addressing health concerns and promoting well-being. Payors can include:

1. Insurance companies provide health insurance plans that cover the cost of medical care for their enrollees.
2. Government agencies providing health insurance coverage for certain populations, such as the elderly, disabled, and low-income individuals.
3. Employers may self-insure their employees' healthcare costs or purchase insurance plans from an insurance company.

Provider: Licensed healthcare providers encompass both individual medical professionals and healthcare organizations that are authorized to offer diagnostic and treatment services related to health conditions. Providers can include:

1. Physicians, also known as doctors, are medical professionals who diagnose and treat diseases and injuries.
2. Nurses provide care to patients in hospitals, clinics, and other healthcare settings.
3. Pharmacists dispense medications and provide counseling on their use.
4. Therapists provide physical, occupational, and speech-language therapy to patients with disabilities or injuries.
5. Hospitals are large healthcare facilities that provide a wide range of medical services, including inpatient care, surgery, and emergency care.

6. Clinics are smaller healthcare facilities that provide outpatient care, such as primary care, dental care, and mental health care.

Health insurance networks, also known as payer networks, play a crucial role in facilitating healthcare coverage. These networks are managed by insurance companies and grant subscribers access to a diverse array of healthcare providers, facilities, and services.

1.4 Existing Problem for Healthcare Payors - Network Creation perspective

The conception of healthcare provider networks was predicated on the assumption that larger networks are invariably superior. This notion is rooted in the belief that an extensive array of providers enhances patient access to healthcare services and could potentially yield improved health outcomes. Nonetheless, empirical studies have indicated that more expansive networks may also contribute to elevated healthcare expenditures. Existing work from. For instance, a study [?] disseminated in 2016 revealed that an augmentation equivalent to one standard deviation in the breadth of a physician network correlated with a **2.8 percent increment in the premiums of healthcare plans**. To put it into perspective, a 10 percent expansion in the physician count within a network could result in an average annual premium hike of **$101**. Author [?] estimated how health insurance premium is associated with physician network breadth, hospital network breadth, and hospital network quality using formula 1

$$log(Premium_{ijcr}) = PhysNetBreadth_i + HospNetBreadth_j + HospNetQuality_j + log(Deductible_i) + log(OOP_i) + PlanType_i + Carrier_c + RatingArea_r + \epsilon_{ijcr}$$

which concludes that cost of designing the healthcare network increase linearly and putting all healthcare provider in the plan is not a feasible solution.

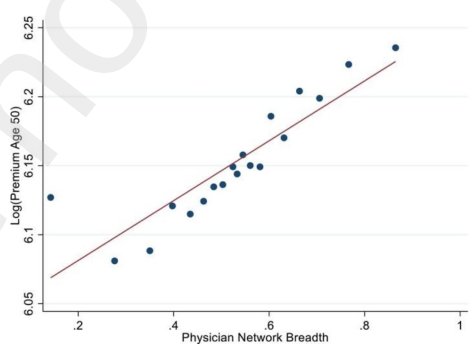

Figure 1: Positive association between provider network breadth and plan premium

1.5 Problem for Healthcare Payors - Adequacy perspective

Healthcare payors face a secondary challenge in that they are not at liberty to autonomously design provider networks; instead, they must adhere to regulatory frameworks established by governments worldwide. These frameworks are in place to guarantee that standard medical criteria are satisfied within each locality. In India, for instance, the National Health Authority (NHA) has set forth a series of directives for the inclusion of healthcare providers into the system. These directives stipulate minimum requirements for staffing, equipment, and facilities that healthcare providers must fulfil to qualify for participation in state- supported health insurance schemes. Similarly, in the United States, the Department of Health and Human Services (HHS) has defined adequacy standards for provider networks within Medicare Advantage and other health insurance programs. These standards aim to ensure that networks possess an adequate quantity and diversity of providers to cater to the healthcare needs of their enrolees. Despite the complexities these regulations introduce, they play a crucial role in ensuring that healthcare provider networks deliver high-quality care and address the healthcare requirements of their members effectively. Collaborating with governmental bodies enables insurance companies to establish networks that not only comply with regulatory mandates but also serve the healthcare needs of their clientele.

For example, if an insurance provider is now coming up with a health plan. It must follow certain guidelines on:

1. Specialist Quota in Facilities
2. Geographic Coverage for Population
3. Limitations on Travel Distance
4. Hospital Bed Availability Ratio
5. Physician Availability Ratio

These requirements play a crucial role in ensuring that healthcare provider networks are comprehensive, accessible, and of adequate quality to meet the needs of their members.

2 Methodology

2.1 Dataset Navigation

The Centre for Medicare and Mediclaim (CMS) maintains a healthcare(USA) dataset that can be used for commercial and academic purposes. CMS website conatins Data Entrepreneurs' Synthetic Public Use File (DE-SynPUF) which is a publicly available dataset that contains synthetic versions of Medicare claims data for the years 2008, 2009, and 2010. The synthetic data is derived from a 5% random sample of Medicare beneficiaries in 2008, with their claims data spanning from 2008 to 2010. Each synthetic beneficiary is assigned a unique unidentifiable ID, DESYNPUFID, which allows for linking synthetic claims to their corresponding synthetic beneficiary.

2.2 Graph & Graph Embeddings

Doctor-patient relationship, geo-graphical locations is mapped to a graph database with each node representing Doctor and Patient with different attributes and edge as the relation between patient and procedure performing provider, and the diagnosis information. Heterogeneous graph is a suitable choice for representing the CMS claims dataset because we have 2 types of Nodes.

- Provider Nodes
- Member Nodes

A Member node can associate with multiple providers, diagnoses, procedures, and medications attributes, whereas a Provider node can be linked to different patients, procedures they perform, and their specialization. A heterogeneous graph can explicitly represent these diverse relationships using different types of Nodes and edges, capturing the intricate connections between entities.
Metapath2vec [**?**] is a popular method for learning node embeddings in graphs. Node embeddings are low-dimensional vector representations that capture the structural and semantic features of nodes in a graph. Node embeddings can be used for various downstream tasks, such as node classification, link prediction, community detection, and graph visualization.The goal of Metapath2vec is to learn a representation for every node based on some combination of its neighboring nodes. The outcome of Metapath2vec node embedding is set of **64-dimensional embeddings** 2. Provider embedding is the representation of attributes like age, gender, ICD-10 [**?**] procedure and diagnostic codes used by the provider. Member embedding is the representation of attributes like birthdate, gender, chronic condition likes existing diabetes, cancer etc.

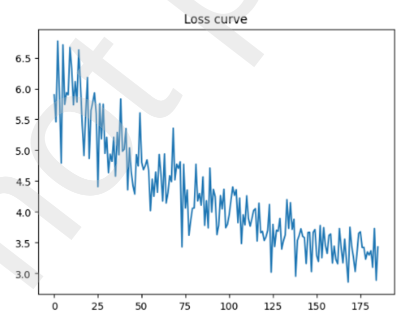

Figure 2: Loss curve for MetaPath2Vec training on CMS graph data

2.3 Network Designing Workflow

The learned node embeddings are used to find the efficiency of the providers by comparing it with the market average3. This efficiency can be used to design the healthcare network with allowing most efficient providers and optimizing the network at the same time by keeping the adequacy similar.Random Forest Regression is an ensemble Machine learning technique used to find the efficiency of the providers.

Training Data: The model is trained on a dataset containing input features.

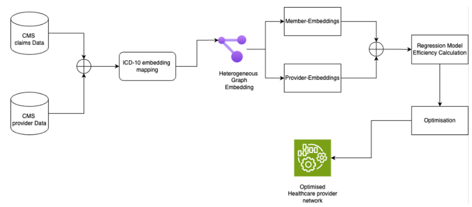

Figure 3: Workflow for healthcare network design

- **feat1** 64 dimensional embedding for provider
- **feat2** 64 dimensional embedding for provider
- **target** variable claim amount

Formulating a mixed-integer programming [**?**] (MIP) model to maximize the efficiency of the provider network while considering constraints related to distance and the number of beds. Let x_i be a binary decision variable that takes the value 1 if provider i is included in the network, and 0 otherwise.

Maximize the sum of the provider efficiency scores for the selected providers using objective function.

$$max \sum_i (ProviderEfficiencyScore_i * x_i)$$

Solving for the network accessibility by adding distance and number of beds constraint while solving the objective function.

3 Results

3.1 Current Findings

By solving this MIP model, you can find the optimal subset of providers that maximizes the overall efficiency score while satisfying the distance and bed capacity constraints.

The Optimization is done on 3 different dimensions.

1. Efficiency: The efficiency of the providers
2. Accessibility: The distance between the geographical centroid of population centers within a state to the geo-coordinates of the provider. This metric tells how far the provider from the population is center.
3. Adequacy: Government norms and adequacy are guardrail on minimum requirements to be met to create any network. For this project we have taken number of beds in the hospital as adequacy check.

On 3 different configuration of networks design 4, we have got a Network-1 with 5768 providers from total 10k providers that are 70% less than the market average in terms of efficiency. With 80% of provider are accessible within 5 Km radius and meeting 95% of minimum bed adequacy. There is observed a trade off between Efficiency and Ac-

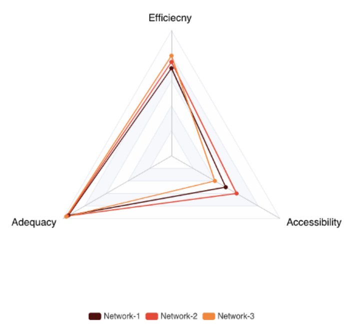

Figure 4: Network Optimization across 3 dimensions

cessibility, overall efficiency of the network reduces if the accessibility is reduced. The system provides quick and efficient network design compared to current handcrafted network, which provides better flexibility and scalability to the insurance Payors.

3.2 Future Scope

- **Integration with Electronic Health Records:** Integrate the graph neural network model with electronic health records (EHRs) to provide real-time recommendations for healthcare network navigation and resource allocation. This could improve care coordination and reduce inefficiencies in the healthcare system.
- **Graph Neural Networks for Healthcare Policy Analysis:** Utilize graph neural networks to simulate the impact of different healthcare policies on network performance and patient outcomes. This could inform evidence-based policymaking and improve the effectiveness of healthcare delivery.