

Лозоватский Илья Маркович
ассистент кафедры компьютерных технологий и
программной инженерии (Кафедра 43)
ФГАОУ ВО «Санкт-Петербургский государственный университет
аэрокосмического приборостроения» (ГУАП)
г. Санкт-Петербург, Россия

Граничное ранжирование графов в транспортных сетях с использованием нейронной сети

Аннотация. Многие сети, такие как транспорт, электроэнергетика и водоснабжение, могут быть представлены в виде графов. Решающей задачей в представлении графов является определение важности (ранжирование) границ графа и их влияния на общую производительность с точки зрения эффективности сети и потока информации. Например, важными границами транспортной сети являются те дороги, воздействие на которые существенно изменит общую эффективность всей сети. Обычно используемый подход к поиску таких важных ребер – это «центрирование ребер между границами» (ЦРГ) – мера ранжирования границ для определения влиятельных ребер графа на основе связности и распространения информации. Вычисление ЦРГ с использованием общего алгоритма Брандеса [1] включает вычисление кратчайших путей для каждой пары узлов, что является достаточно трудоемким, дорогостоящим и ограниченным в вычислительном отношении методом, особенно для больших графов. Любые изменения параметров графа, например веса ребра или добавление и удаление узлов или ребер, требуют пересчета ЦРГ. В качестве основного нововведения в данной статье предложен приближенный метод оценки ЦРГ с использованием нейронной сети графов (НСГ), подхода, основанного на глубоком обучении. Аргументировано положение о том, что он более эффективен в вычислительном отношении по сравнению с традиционным методом, особенно для больших графов. Предлагаемый метод ранжирования ребер на основе НСГ оценивается на нескольких синтезированных графах и реальном наборе транспортных данных. Показано, что эта структура может оценить приблизительный рейтинг ребер намного быстрее по сравнению с традиционным методом, предложенным Брандесом [1]. Этот подход является индуктивным, т. е. обучение и тестирование выполняются на разных наборах графов с различным количеством узлов и ребер. Предлагаемый метод особенно подходит для приложений в крупномасштабных сетях, где требуется периферийная информация, например, в проектах улучшения городской инфраструктуры, анализе устойчивости сетей электро- и водоснабжения, а также оптимизации распределения ресурсов в инженерных сетях.

Ключевые слова: транспортная сеть, распределение ресурсов, ранжирование ребер графа, нейронная сеть, глубокое обучение

Lozovatsky Ilya Markovich
assistant of the department of computer technology and
software engineering (department 43)
of the St. Petersburg State University of Aerospace Instrumentation
St. Petersburg, Russia

Boundary ranking of graphs in transport networks using a neural network

Annotation. Many networks, such as transport, electric power and water supply, can be represented as graphs. The crucial task in graph representation is to determine the importance (ranking) of graph boundaries and their impact on overall performance in terms of network efficiency and information flow. For example, the important boundaries of the transport network are those roads, the impact on which will significantly change the overall efficiency of the entire network. A commonly used approach to finding such important edges is "edge-to-edge centering" (CRG), a measure of boundary ranking to identify influential edges of a graph based on connectivity and information dissemination. Calculating the CRG using the general Brandes algorithm [1] involves calculating the shortest paths for each pair of nodes, which is a rather time-consuming, expensive and computationally limited method, especially for large graphs. Any changes to the graph parameters, such as edge weights or the addition and removal of nodes or edges, require recalculation of the CRG. As the main innovation in this article, an approximate method for estimating the CRG using a neural network of graphs (NSG), an approach based on deep learning, is proposed. It is argued that it is more computationally efficient than the traditional method, especially for large graphs. The proposed method of edge ranking based on NG is evaluated on several synthesized graphs and a real set of transport data. It is shown that this structure can estimate the approximate edge rating much faster than the traditional method proposed by Brandes [1]. This approach is inductive, i.e. training and testing are performed on different sets of graphs with different numbers of nodes and edges. The proposed method is particularly suitable for applications in large-scale networks where peripheral information is required, for example, in urban infrastructure improvement projects, analysis of the stability of electricity and water supply networks, as well as optimization of resource allocation in utility networks.

Keywords: transport network, resource allocation, graph edge ranking, neural network, deep learning

Введение

Транспортные сети играют решающую роль в экономике и благосостоянии граждан, обеспечивая беспрепятственное перемещение людей и товаров, а также выступая в качестве «артерий» для эвакуации во время катастроф и стихийных бедствий. Здоровая транспортная сеть предлагает значительные преимущества своим гражданам, например, за счет мобильности капитала и рабочей силы, диффузии населения или национальной обороны [2]. Стихийные бедствия и

сезонные проявления погоды могут серьезно влиять на транспортные сети, что приведет к прямым потерям, таким как затраты на ремонт инфраструктуры, и косвенным потерям, таким как снижение эффективности сети [3]. В то время как некоторые элементы системы автомагистралей ремонтируются и возобновляются в течение нескольких недель, другие элементы могут оставаться непроходимыми в течение многих месяцев. Понимание производительности, пропускной способности и критических участков дорог имеет важное значение для принятия решений во время таких катастрофических событий. Одной из важнейших задач, связанных с таким пониманием, является объективное определение того, какие сегменты дорог имеют решающее значение для эффективности системы в целом.

Ответом на этот вопрос является рассмотрение сети как графа, состоящего из узлов и ребер, представляющих участки дорожных развязок и сегментов соответственно, и оптимизация ключевых показателей, таких как стоимость времени в пути или расстояние, с использованием этого графа путем привязки весов к ребрам. Такие узлы или ребра можно удалить или изменить веса ребер, чтобы имитировать потерю или снижение функциональности в локальных регионах сети. Ранжирование узлов обеспечивает естественное применение в социально-значимых сетях [4]; с другой стороны, ранжирование важности границ больше подходит для инженерных сетевых систем, таких как транспортные сети. Выявление критических границ в транспортных сетях может повлиять на упреждающие стратегии устранения недостатков в основных сегментах системы, тем самым делая всю систему более надежной и устойчивой к сбоям. Основная цель настоящей статьи – предложить новый и эффективный в вычислительном отношении способ оценки важности ребер в транспортных сетях. Показано, что этот приближенный метод, основанный на НСГ, превосходит традиционные методы с точки зрения скорости, достигая при этом сопоставимого уровня производительности.

Методология

Модели нейронных сетей для данных, структурированных графами, известны как НСГ [5,6]. Эти модели используют структуру графа для агрегирования информации о характеристиках/вложениях ребер и узлов [6]. Агрегирование признаков из структурированного шаблона графика позволяет НСГ предсказывать вероятность существования ребер или предсказывать метки узлов.

Информация о структуре графа усваивается из матрицы смежности и матрицы информации о признаках узлов и ребер, которые формируют входные данные, и обучается с использованием функции потерь. Передача сообщений происходит на каждом уровне НСГ, когда каждый узел агрегирует признаки своих соседей. Вектор объектов узла обновляется путем объединения его вектора объектов с агрегированными объектами из соседних узлов. В первом слое GNN объединяет признаки своих непосредственных соседей, и с увеличением числа слоев глубина усвоения признаков соседних краев соответственно увеличивается. Вектор краевых объектов обновляется с помощью

агрегированных объектов с соседних краев, и эта процедура повторяется для каждого слоя НСГ. В НСГ есть три этапа, разработанные следующим образом:

Шаг 1: Передача сообщений о пограничных функциях:

Оригинальный алгоритм передачи сообщений применяется к узловым объектам, проходящим по ребрам графа. Поскольку настоящая статья сосредоточена на ранжировании ребер, каждое ребро связано с граничными объектами/встраиванием. Такие объекты представлены с использованием векторного R^d – скрытого размерного представления. Популярным алгоритмом для получения такого представления является Node2vec [7].

В частности, Node2vec использует пропускную грамму (skip-gram) с отрицательной выборкой. Основная идея skip-gram заключается в максимизации вероятности предсказания правильного контекстного узла с учетом центрального узла. Процесс skip-gram для встраивания узла показан на рисунке 1.

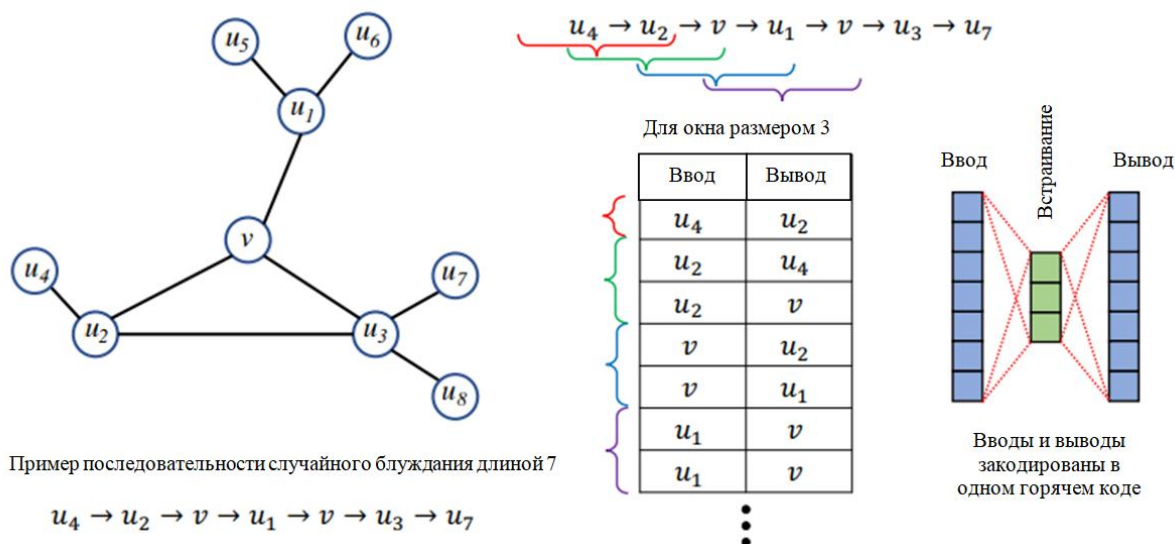


Рисунок 1– Иллюстрация модели skip-gram для встраивания узлов.

(Для простого случайного блуждания длиной 7 скользящее окно длиной 3 используется для подготовки входных и выходных данных для обучения модели Word2Vec.)

Следовательно, в статье используется быстрая и экономичная по памяти версия Node2Vec, называемая PscanPy (Распараллеленное, экономичное по памяти и ускоренное Node2Vec в Python) [8]. PscanPy делает реализацию Node2Vec эффективной по следующим трем направлениям:

Параллелизм: Оценка вероятности перехода и процессы генерации случайных блужданий независимы, но не распараллелены в оригинальном Node2Vec. PscanPy распараллеливает процесс генерации обхода, что значительно ускоряет операцию.

Структура данных: Исходная реализация Node2Vec использует NetworkX [9] для хранения графиков, что неэффективно для крупномасштабных вычислений. Однако PscanPy использует формат Compact Sparse Row (CSR) для разреженных графиков, который имеет свойства разреженности, аналогичные транспортной сети, рассматриваемой в настоящей статье. Графики в формате CSR более компактно хранятся в памяти и работают быстрее, поскольку могут

более эффективно использовать кеш.

Память: Исходная версия Node2Vec предварительно обрабатывает и сохраняет вероятности перехода 2-го порядка, что приводит к значительному использованию памяти. ResanRu исключает этап предварительной обработки и вычисляет вероятности перехода, когда это необходимо, без сохранения.

Новый фреймворк, представленный в настоящей статье, содержит модифицированную версию концепции передачи сообщений – функции ребер агрегируются и передаются соседним ребрам. Таким образом, НСГ изучает структурную информацию. Пример шага передачи сообщений показан на рисунке 2. В то время как обычная реализация НСГ использует передачу сообщений по узлам, здесь такая передача выполняется по границам, что является новшеством.

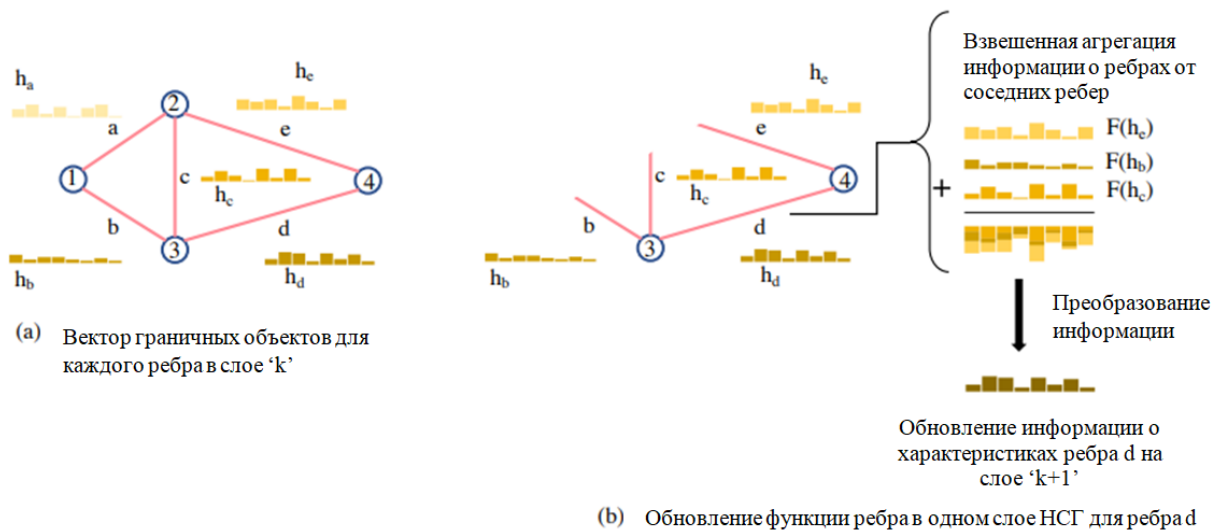


Рисунок 2– Принципиальная схема передачи сообщений НСГ: (а) Это пример графика с 4 узлами и 5 ребрами, (б) процедура передачи сообщения ребру d с использованием векторов ребер соседних ребер

Шаг 2: Агрегирование:

Сообщения агрегируются после того, как все сообщения с соседних ребер будут переданы интересующему ребру.

Некоторыми популярными функциями агрегирования являются:

$$\begin{aligned} Sum &= \sum_{j \in N_i} F(h_j); & Mean &= \frac{\sum_{j \in N_i} F(h_j)}{|N_i|} \\ Max &= \max_{j \in N_i} F(h_j); & Min &= \min_{j \in N_i} F(h_j) \end{aligned} \quad (1)$$

где N_i - множество соседних ребер ребра i (представляющего интерес ребра). Рассматривая "AGGREGATE" как функцию агрегирования (сумма, среднее значение, максимальное или минимальное значение преобразования сообщения соседнего ребра), агрегированное сообщение μ на уровне k может быть выражено как:

$$\mu_i^{(k)} = AGGREGATE^{(k)}(\{h_j^{(k)} : j \in N_i\}) \quad (2)$$

Шаг 3: Обновление:

Эти агрегированные сообщения обновляют объекты исходного ребра на уровне GNN. На этом обновленном шаге вектор объектов ребра объединяется с агрегированными сообщениями и выполняется простым добавлением или объединением:

$$\text{Добавление: } h_j^{(k+1)} = \sigma(\Gamma(\Omega(h_j^{(k)} + \mu_j^{(k)}))); \quad (3)$$

$$\text{Объединение: } h_j^{(k+1)} = \sigma(\Gamma(\Omega(h_j^{(k)} \oplus \mu_j^{(k)})))$$

где σ - функция активации, Ω - простой многослойный перцептрон (ПМП), а Γ - другая нейронная сеть, которая проецирует добавленные или объединенные векторы в другое измерение. Вкратце, шаг обновления с предыдущего уровня можно резюмировать следующим образом:

$$h_j^{(k+1)} = COMBINE^{(k)}(h_j^{(k)}, \mu_j^{(k)}) \quad (4)$$

Выходные данные каждого уровня НСГ передаются в качестве входных данных на следующий уровень НСГ. После k -го NN слоев/итерации вектор встраивания ребер на конечном слое захватывает информацию об особенностях ребер и информацию о структуре графа всех смежных ребер от расстояния в 1 переход до k -го расстояния перехода. Вектор краевых признаков 1-го слоя получен с помощью NodeVec [7]

Предлагаемая структура НСГ

Основываясь на концепциях НСГ, представленных ранее, предложена авторская структура НСГ для оценки приблизительного ранжирования по границам. Здесь основная причина выбора НСГ вместо ССГ (Сеть свертки графов) [10] заключается в том, что в агрегированном векторе признаков конкретного ребра доминирует вектор признаков самого ребра, поскольку ССГ добавляет самоповтор в граф. Далее представлен алгоритм вместе с предлагаемой архитектурой. Приводится описание модифицированной матрицы смежности и ее использования для ранжирования границ между объектами.

Алгоритм и архитектура НСГ

На рисунке 3 показан общий процесс вычисления приблизительного ЕВС. Эта структура использует структуру графа – в частности, матрицу смежности ребер – и матрицу признаков в качестве входных данных для оценки вектора ранжирования ЕВС, отражающего важность каждого ребра в структуре графа. В основе этой процедуры лежит модуль GNN, входными данными которого являются матрица краевых признаков и два варианта матрицы смежности краев. Начиная с начальных весов в модуле GNN, вектор ранжирования ЕВС модели вычисляется путем обратного распространения ошибок по слоям GNN и последующего итеративного обновления весов.

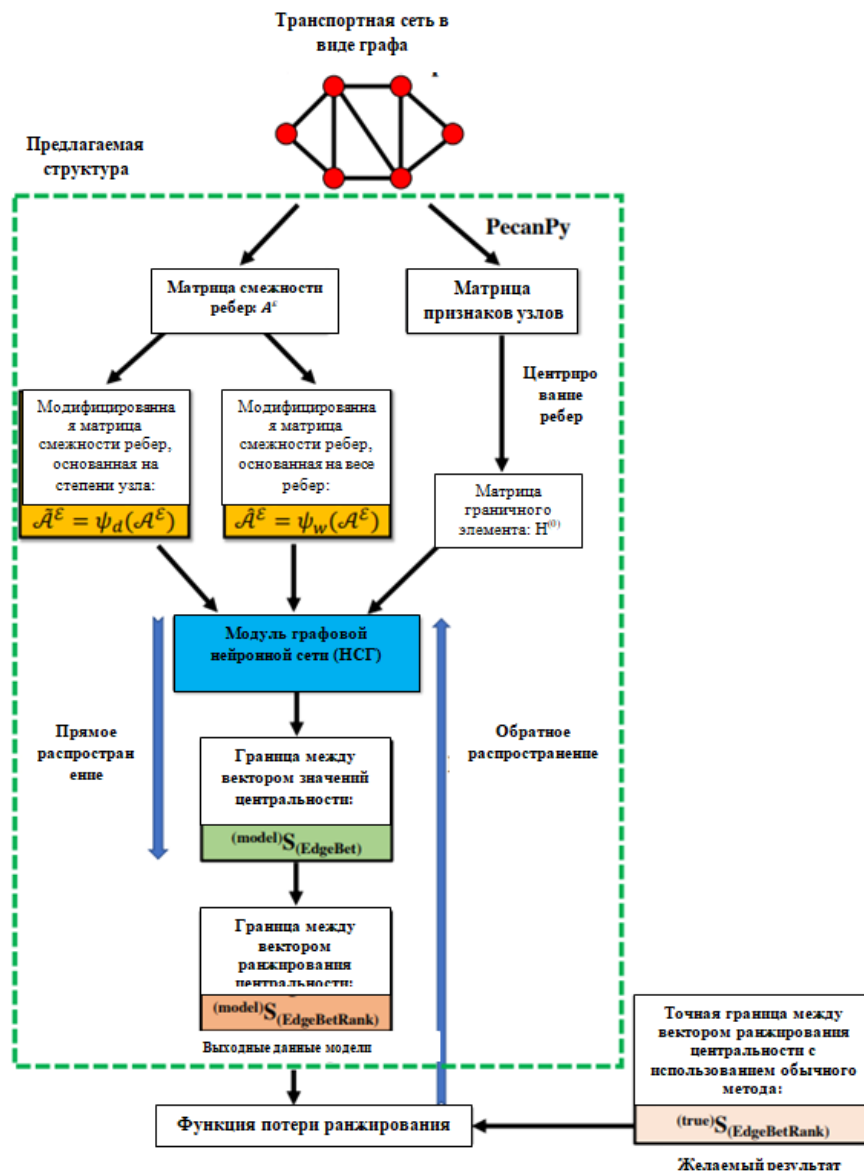
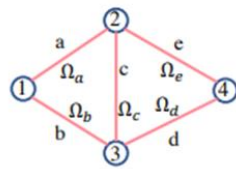


Рисунок 3– Предлагаемая структура для ранжирования центрирования ребер между границами

Матрица смежности ребер

Матрица смежности ребер не уникальна для всех графовых структур. Например, пара неизоморфных графов - трехточечные звездообразные графы S_3 и циклический граф с тремя вершинами C_3 имеют идентичные матрицы смежности ребер, как показано на рисунках 4. Следовательно, вводятся два варианта для этой матрицы – модифицированную матрицу смежности ребер, основанную на степени узла A^E , и модифицированную матрицу смежности ребер, основанную на весе ребра A^E . Модифицированные матрицы смежности ребер получены из матрицы смежности ребер с использованием функций ψ_d и ψ_w соответственно – подробно показано в алгоритме 1: строки 13-32, а соответствующий пример показан на рисунке 3. Веса ребер a, b, c, d и e равны $\Omega_a, \Omega_b, \Omega_c, \Omega_d$ и Ω_e соответственно. Матрица A^E уникальна для каждого графа; A^E сохраняет сходные свойства с исходной матрицей смежности ребер и не является уникальной для графовых структур.



Пример графа

	a	b	c	d	e
a	0	1	1	0	1
b	1	0	1	1	0
c	1	1	0	1	1
d	0	1	1	0	1
e	1	0	1	1	0

\mathcal{A}^E

Матрица смежности ребер

	a	b	c	d	e
a	0	1/2	1/3	0	1/3
b	1/2	0	1/3	1/3	0
c	1/3	1/3	0	1/3	1/3
d	0	1/3	1/3	0	1/2
e	1/3	0	1/3	1/2	0

$\bar{\mathcal{A}}^E$

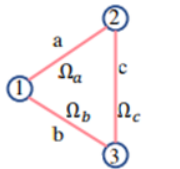
Модифицированная матрица смежности ребер, основанная на степени узла

	a	b	c	d	e
a	0	$2/(\Omega_a + \Omega_b)$	$2/(\Omega_a + \Omega_c)$	0	$2/(\Omega_a + \Omega_e)$
b	$2/(\Omega_a + \Omega_b)$	0	$2/(\Omega_b + \Omega_c)$	$2/(\Omega_b + \Omega_d)$	0
c	$2/(\Omega_a + \Omega_c)$	$2/(\Omega_b + \Omega_c)$	0	$2/(\Omega_c + \Omega_d)$	$2/(\Omega_c + \Omega_e)$
d	0	$2/(\Omega_b + \Omega_d)$	$2/(\Omega_c + \Omega_d)$	0	$2/(\Omega_d + \Omega_e)$
e	$2/(\Omega_a + \Omega_e)$	0	$2/(\Omega_c + \Omega_e)$	$2/(\Omega_d + \Omega_e)$	0

$\bar{\mathcal{A}}^E$

Модифицированная матрица смежности ребер, основанная на весе ребер

Рисунок 3– Модифицированные матрицы смежности ребер из матрицы смежности ребер выборочной графовой сети

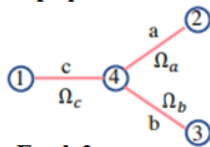


Граф-1

	a	b	c
a	0	1	1
b	1	0	1
c	1	1	0

	a	b	c
a	0	1/2	1/2
b	1/2	0	1/2
c	1/2	1/2	0

	a	b	c
a	0	$2/(\Omega_a + \Omega_b)$	$2/(\Omega_a + \Omega_c)$
b	$2/(\Omega_a + \Omega_b)$	0	$2/(\Omega_a + \Omega_c)$
c	$2/(\Omega_a + \Omega_c)$	$2/(\Omega_a + \Omega_c)$	0



Граф-2

	a	b	c
a	0	1	1
b	1	0	1
c	1	1	0

\mathcal{A}^E

Матрица смежности ребер

	a	b	c
a	0	1/3	1/3
b	1/3	0	1/3
c	1/3	1/3	0

$\bar{\mathcal{A}}^E$

Модифицированная матрица смежности ребер, основанная на степени узла

	a	b	c
a	0	$2/(\Omega_a + \Omega_b)$	$2/(\Omega_a + \Omega_c)$
b	$2/(\Omega_a + \Omega_b)$	0	$2/(\Omega_a + \Omega_c)$
c	$2/(\Omega_a + \Omega_c)$	$2/(\Omega_a + \Omega_c)$	0

$\bar{\mathcal{A}}^E$

Модифицированная матрица смежности ребер, основанная на весе ребер

Рисунок 4 – Неуникальность матрицы смежности ребер и соответствующие модификации, предложенные в новой структуре
Функция потерь

Функция потери рейтинга используется для оценки потерь из-за различий в рейтинге, прогнозируемом предлагаемой моделью, по сравнению с целевым рейтингом ГМЦ. Такие функции потери рейтинга ранее использовались для систем рекомендаций для ранжирования или оценки продуктов или пользователей [11]. Функция потерь при ранжировании маржи определяется следующим образом:

$$\mathcal{L}(S_i^{(model)}, S_i^{(true)}, y) = \max(0, -y \cdot (S_i^{(model)} - S_i^{(true)})) + Margin \quad (5)$$

$$y = \begin{cases} 1, & \text{если } S_i^{(model)} \text{ должен быть оценен выше, чем } S_i^{(true)} \\ -1, & \text{если } S_i^{(true)} \text{ должен быть оценен выше, чем } S_i^{(model)} \end{cases}$$

где $S_i^{(model)}$, - прогнозируемая оценка ранжирования, а $S_i^{(true)}$ - оценка ГМЦ, полученная с использованием обычного метода, т.е. алгоритма Брандеса [12]. В этом исследовании значение резерва установлено равным 1, чтобы обеспечить некоторую гибкость.

Результаты

Параметры генерации графика

Параметры генерации синтезированного графика приведены в таблице 1. Здесь $U\{a, b\}$ и $U[a, b]$ представляют дискретные и непрерывные равномерные распределения между диапазонами a и b соответственно. С параметрами генерации графиков, выбранными произвольно для этого тематического исследования, генерируются 1000 обучающих графиков, 100 графиков проверки и 100 тестовых графиков для каждого случая.

Таблица 1– Параметры генерации синтезированных графов

Тип синтезированного графа	Параметры генерации	
Модель Эрдёша – Реньи I (GNP)	Номера узлов Вероятность создания ребра Веса ребер	$U\{1000,5000\}$ $1.2/(\text{Номер узла} - 1)$ $U\{0,100\}$
Модель Эрдёша – Реньи II (GNM)	Номера узлов Номера ребер Веса ребер	$U\{1000,5000\}$ $U[1.4,1.6] \times$ Номер узла $U[0,100]$
Модель Уоттса – Строгаца	Номера узлов Средняя степень Вероятность перемонтажа ребер Веса ребер	$U\{2000,4000\}$ 4 0.5 $U[0,100]$

В таблице 2 обобщены результаты, полученные на основе синтезированных графиков. Средняя длина кратчайшего пути определяется как:

$$a = \sum_{s,t \in V} \frac{d(s,t)}{|V| \cdot (|V|-1)} \quad (6)$$

где V - множество узлов во взвешенном/невзвешенном графе G общего числа узлов $|V|$ (мощность), а $d(s, t)$ - кратчайший путь от s до t . Коэффициент кластеризации - это мера для узлов, которые сгруппированы вместе, которая представляет собой среднее геометрическое значение весов ребер подграфа [13].:

$$c_s = \frac{\sum_{tu} (\hat{\omega}_{st} \hat{\omega}_{tu} \hat{\omega}_{us})}{\deg(s)(\deg(s)-1)}; \quad C = \frac{1}{|V|} \sum_{s \in V} c_s \quad (7)$$

где cs - коэффициент кластеризации узлов, C - средний коэффициент кластеризации графа; $\text{deg}(s)$ - степень узла s ; узлы s , t и u создают треугольники на графике. Веса ребер $\hat{\omega}_{st}$ нормализуются на максимальный вес в сети, $\hat{\omega}_{st} = \omega_{st}/\max(\omega)$.

Таблица 2– Статистика синтезированных графиков, используемых для обучения и тестирования предлагаемой платформы НСГ

Тип синтезированного графа	Модель Эрдёша – Реньи I (случайный выбор)	Модель Эрдёша – Реньи II (случайный выбор)	Сеть «малого мира» Модель Уоттса–Строгаца
Количество узлов (Диапазон)			
Количество ребер (Диапазон)	До 5000	До 5000	До 5000
Средняя длина кратчайшего пути ($\mu \pm \sigma$, max)	277.5 ± 15.7, 307.5	290.6 ± 20.6, 338.6	249.4 ± 7.4, 266.6
Средний коэффициент кластеризации ($\mu \pm \sigma$)	9.04 x 10 ⁻⁴ ±7.91 x 10 ⁻⁴	1.02 x 10 ⁻⁴ ±7.97 x 10 ⁻⁴	0.0681 ±0.004
Средняя степень образования узлов ($\mu \pm \sigma$)	3.198 ±0.031	3.182 ±0.102	4 ± 0

Приложение к транспортной сети Краснодарского края

Структура была проверена на основе НСГ на транспортной сети Краснодарского края. Сетевая информация получается из сетевого репозитория, который содержит 2640 узлов (дорожных развязок) и 3302 ребер (улиц). Эта сеть иллюстрирует хороший баланс размера и вычислительных затрат для демонстрации аспектов производительности предлагаемой структуры. Рисунок 5 показывает оценки ГМЦ для сети. На этом рисунке улицы, выделенные красным, обозначают наиболее важные дороги, на графике, как это определено мерой ГМЦ – любое изменение этих путей (изменение весов ребер, добавление и удаление дорог и перекрестков) будет существенно влиять на сеть. Это исследование моделирует ранжирование важности границ из-за динамического изменения параметров сети, например, строительства новых дорог или выявление неработающих/перекрытых дорог, или изменения других факторов, таких как время в пути, расстояние поездки или транспортный поток.

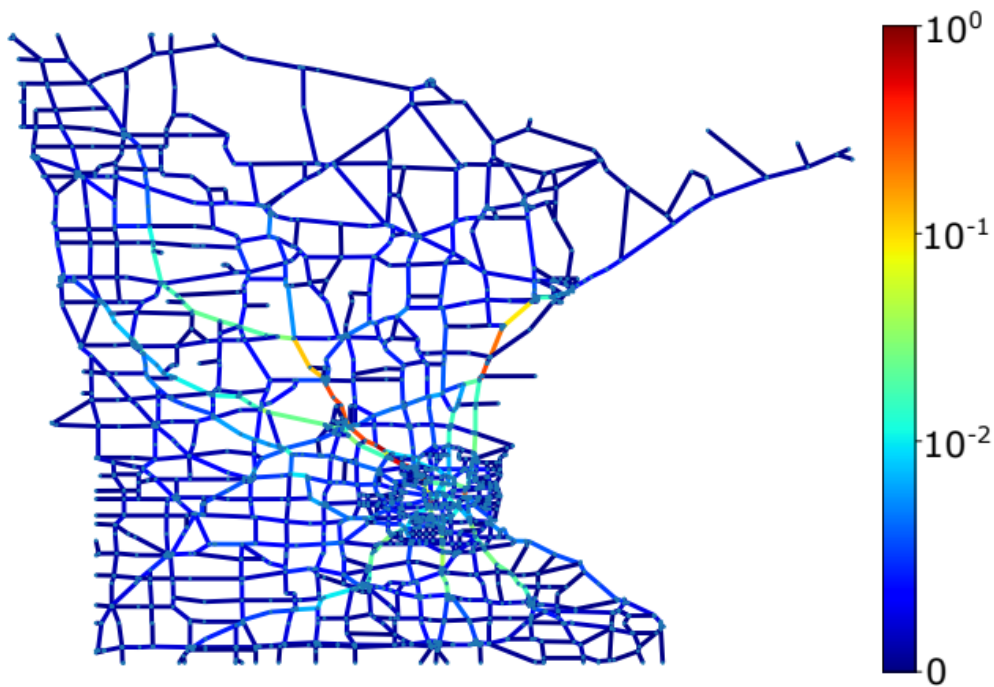


Рисунок 5– Разница в показателях центральности транспортной сети КК

В экспериментах рассматриваются два случая: (а) изменение весов ребер, например, возникающее из-за изменения объемов трафика, и (б) изменение весов ребер в сочетании с постоянным изменением (удалением или добавлением) ребер, скажем, из-за различных объемов трафика в сочетании с некоторыми неработоспособными дорогами или добавлением новых дорог. В обоих этих случаях веса ребер определяются исходя из евклидова расстояния (2-норма) между координатами узлов, т.е. длины сегмента дороги. Эти случаи имитируют посткатастрофические события, такие как сильное оползни, обрушение, наводнение и пр., где первый случай может быть предшествующим событию или после незначительных событий, когда все пути работают, но тем не менее, их свойства изменились. Последнее может быть результатом значительного события, когда некоторые сегменты выводятся из эксплуатации, а остальные функционируют со значительно измененными свойствами.

Пример I: Изменение весов ребер

В этом случае веса ребер изменяются случайным образом в соответствии с $r \times \omega_i$, где ω_i - вес ребра i , а значения r выбираются из непрерывного равномерного распределения $U[0.8, 1.2]$. Количество узлов и ребер для графиков остается неизменным. Статистика графиков приведена в таблице 3. Из этой модели генерации графиков было выбрано 1000 обучающих данных и 200 тестовых данных, а оценки показаны в таблице 4 и на рисунке 6. Что касается синтетических графиков, то высокие баллы и относительно небольшое стандартное отклонение указывают на то, что модель достаточно надежна.

Таблица 3– Статистические данные исследования транспортной сети (пример I)

Количество узлов	Количество ребер	Средняя длина кратчайшего пути ($\mu \pm \sigma$, max)	Средний коэффициент кластеризации. ($\mu \pm \sigma$)	Средняя степень образования узлов ($\mu \pm \sigma$)
640	302	146.5 \pm 0.72, 148.7	0.016 \pm 6.0 x 10 ⁻¹⁸	2.501 \pm 8.88 x 10 ⁻¹⁶

Таблица 4– Рейтинговые оценки по обучению и тестированию для транспортной сети (пример I)

	Обучающие данные	Тестовые данные
Оценка Кендалла Tau ($\mu \pm \sigma$)	0.747 \pm 0.010	0.694 \pm 0.014
Коэффициент корреляции Спирмена ($\mu \pm \sigma$)	0.911 \pm 0.008	0.865 \pm 0.013

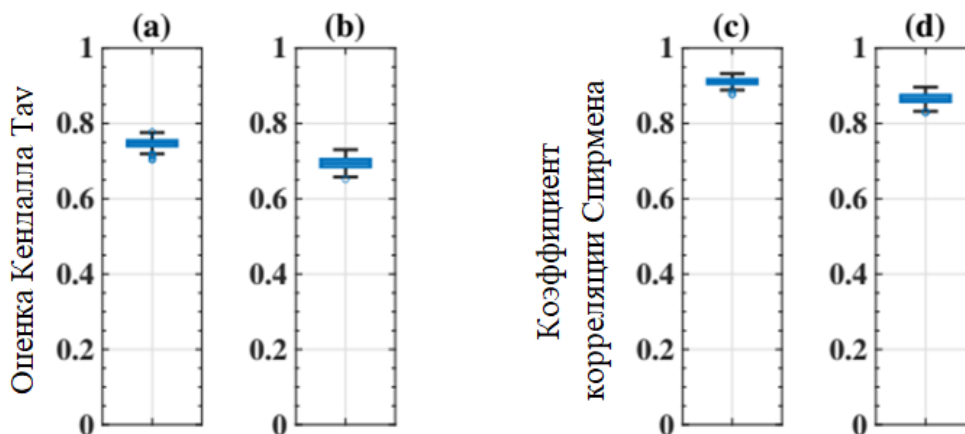


Рисунок 6 – Распределение рейтинговых баллов по обучению и тестированию для транспортной сети (пример I). [График прямоугольника и усов показывает медиану, нижний и верхний квартили, любые выбросы (рассчитанные с использованием межквартильного диапазона), а также минимальные и максимальные значения, которые не являются выбросами] – (а) оценка Кендалла-тау по данным обучения, (б) оценка Кендалла-Тау по данным тестирования, (с) коэффициент корреляции Спирмена по данным обучения, (д) коэффициент корреляции Спирмена по данным тестирования.

Пример II: Одновременное изменение веса ребер и количества ребер

В этом случае веса ребер такие, как указано в Примере 1, т.е. $r \times \omega_i$ с $r \in U[0.8, 1.2]$. Кроме того, изменено количество ребер в графе. Количество ребер выбирается из дискретного равномерного распределения $U\{3269, 3302\}$, которое основано на удалении максимум 1% ребер из исходной сети. Рассмотренный случай, т.е. удаление границ на 1%, отражает экстремальный случай, следующий за катастрофическим стихийным событием. Это моделирование также включает

добавление новых дорог. Статистические данные графика приведены в таблице 5. После обучения с использованием 1000 тренировочных данных оценка оценивается по 200 тестовым данным и тренировочным данным и показана в таблице 6 и на рисунке 7. Как и в случае I, стандартное отклонение рейтинговой оценки относительно невелико, что указывает на надежность предлагаемого метода.

Таблица 5– Статистические данные исследования транспортной сети (пример II)

номер узла	номер ребра	Средняя длина кратчайшего пути ($\mu \pm \sigma$, max)	Средний коэффициент кластеризации. ($\mu \pm \sigma$)	Средняя степень образования узлов ($\mu \pm \sigma$)
640	269-3302	147.6 \pm 1.08, 151.4	0.016 \pm 3.2 x 10 ⁻⁴	2.49 \pm 7.2 x 10 ⁻³

Таблица 6– Рейтинговые оценки по обучению и тестированию для транспортной сети (пример II)

Оценка ($\mu \pm \sigma$)	Кендалла	Тау	Обучающие данные	Тестовые данные
Кoeffициент корреляции Спирмена ($\mu \pm \sigma$)			0.708 \pm 0.019	0.63 \pm 0.023
			0.881 \pm 0.015	0.808 \pm 0.022

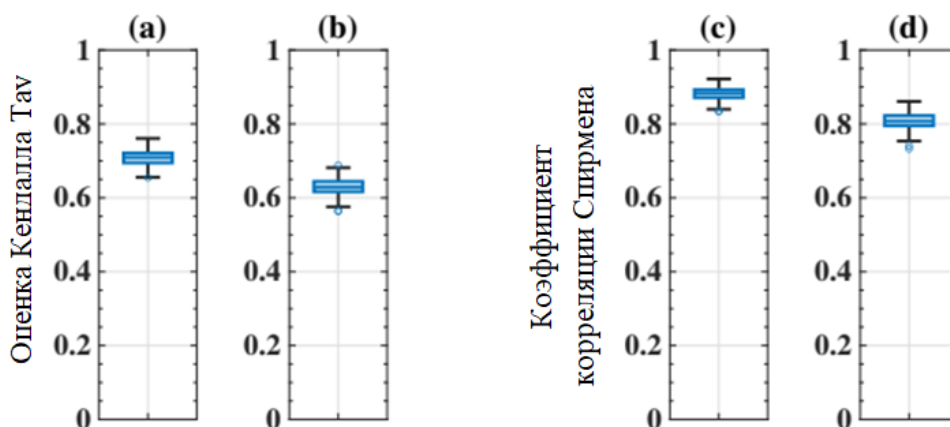


Рисунок 7 – Распределение рейтинговых баллов по обучению и тестированию для транспортной сети Краснодарского края (пример II). График прямоугольника и усов показывает медиану, нижний и верхний квартили, любые выбросы (рассчитанные с использованием межквартильного диапазона), а также минимальные и максимальные значения, которые не являются выбросами – (а) балл Кендалла-тау по данным обучения, (б) балл Кендалла-тау по данным тестирования, (в) Кoeffициент корреляции Спирмена по данным

обучения, (d) коэффициент корреляции Спирмена по данным тестирования.

Выводы

Оценка важности ребер в графе имеет решающее значение во многих приложениях, например, в транспортных, энергетических и водопроводных сетях. Традиционный подход, основанный на кратчайшем пути, может решить проблему за счет значительных вычислительных затрат. Было показано, что использование нейронных сетей графов (GNN) для аппроксимации важности ребер может решить эту проблему лишь за счет умеренного снижения точности. Такие сокращения могут быть приемлемыми в чрезвычайных ситуациях, когда задержка является основной проблемой. Время вычислений предлагаемого метода имеет линейный порядок, что превосходит сложность квадратичного порядка традиционного метода. Результаты модели НСГ, с точки зрения ранжирования ребер, хорошо коррелируют со значениями ГМЦ, полученными с использованием традиционного подхода. Эта приблизительная оценка ранжирования ребер выполняется быстрее для больших графовых сетей, скажем, с числом узлов более 2000, что является основным преимуществом.

Предложенная модель НСГ является масштабируемой: количество ребер в графе может варьироваться, и модель можно применять к графам разных размеров. Она также обобщает, поскольку проверка выполняется на данных, невидимых моделью на этапе обучения. Предложенная структура проверена на трех типах синтетических графов: Эрдеша-Реньи-I, Эрдеша-Реньи-II и модели Уоттса-Строгаца. Средний коэффициент корреляции Спирмена для всех этих случаев в тестовом наборе данных составляет примерно 0,92, что указывает на эффективность предлагаемого метода. Также была изучена эффективность предлагаемой структуры в транспортной сети Краснодарского края. Для этого было рассмотрено два случая – (а) изменение веса ребра и (б) изменение веса ребра вместе с удалением/добавлением ребра. Для этих случаев коэффициенты корреляции Спирмена составлены 0,865 и 0,808 соответственно. Результаты показывают, что предлагаемый метод имеет отличный потенциал.

Список источников

1. U. Brandes, A faster algorithm for betweenness centrality, *Journal of mathematical sociology* 25 (2) (2001) 163–177
2. N. Nadh, Important functions of transportation network (2022).
3. B. Postance, J. Hillier, T. Dijkstra, N. Dixon, Extending natural hazard impacts: an assessment of landslide disruptions on a national road transportation network, *Environmental Research Letters* 12 (1) (2017)
4. A. Zareie, A. Sheikahmadi, M. Jalili, Influential node ranking in social networks based on neighborhood diversity, *Future Generation Computer Systems* 94 (2019) 120–129
5. Павлов Д.А., Кочкаров А.А. Об одной многокритериальной задаче покрытия минимального веса предфрактального графа простыми пересекающимися цепями. Препринт №200. РАН САО . Нижний Архыз. 2004.-

12с.

6. Кочкаров А. А., Кочкаров Р. А. Параллельный алгоритм поиска кратчайшего пути на предфрактальном графе / А. А. Кочкаров, Р. А. Кочкаров // Журн. вычислит. матем. и матем. физики. 2004. Т. 44. № 6. С. 1157–1162.

7. A. Grover, J. Leskovec, node2vec: Scalable feature learning for networks, in: Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, 2016, pp. 855–864

8. R. Liu, A. Krishnan, Pecanpy: a fast, efficient and parallelized python implementation of node2vec, Bioinformatics 37 (19) (2021)

9. Кочкаров А. М. Структурная динамика и исследование структурновременных характеристик дискретных систем / А. М. Кочкаров, А. А. Кочкаров, С. П. Никищенко // Известия ТРТУ. Тематический выпуск “Перспективные системы и задачи управления”. – Таганрог: ТРТУ, 2006. – № 3. – С. 235-238.

10. Малинецкий Г. Г. Математические основы синергетики. Хаос, структуры, вычислительный эксперимент. М.: КомКнига, 2005.

11. C. J. Burges, From ranknet to lambdarank to lambdamart: An overview, Learning 11 (23-581) (2010)

12. У. Брандес, О вариантах кратчайшего пути между центральностью и их общим вычислением, Социальные сети 30 (2) (2008)

13. Урыков В.А., Зеленина Л.И. Математические модели транспортных потоков // Современная техника и технологии. 2015. № 6 [Электронный ресурс]. URL: <http://technology.snauka.ru/2015/06/6051> (дата обращения: 17.12.2023).