

РАЗРАБОТКА АРХИТЕКТУР ДИНАМИЧЕСКИХ GNN ДЛЯ МОДЕЛИРОВАНИЯ ЭВОЛЮЦИИ СОЦИАЛЬНЫХ СЕТЕЙ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ОБРАБОТКИ ВРЕМЕННЫХ ГРАФОВ ИЗМЕНЕНИЙ СТРУКТУРЫ СОЦИАЛЬНЫХ СЕТЕЙ В РЕАЛЬНОМ ВРЕМЕНИ

Попов Валерий Владиславович
Аспирант, Российский Технологический
Университет МИРЭА
popov.v.v.v.v.p@gmail.com

DEVELOPMENT OF DYNAMIC GNN ARCHITECTURES FOR MODELING THE EVOLUTION OF SOCIAL NETWORKS: A COMPARATIVE ANALYSIS OF METHODS FOR PROCESSING TEMPORAL GRAPHS OF STRUCTURAL CHANGES IN SOCIAL NETWORKS IN REAL TIME

V. Popov

Summary. Social networks represent complex dynamic systems where connections between users and entities (posts, comments, communities) change in real time. Modeling such temporal structures for further analysis using classical mathematical modeling methods is challenging due to high dimensionality and the need to account for data additivity. Traditional graph neural networks (GNN — Graph Neural Network) demonstrate high efficiency in analyzing static graphs; however, their application is limited when working with temporal data that requires adaptation to changes in graph structure. This article examines modern approaches to the development of dynamic GNNs capable of modeling changes in social networks while considering temporal patterns. An architecture combining recurrent mechanisms and neighbor aggregation is proposed for processing streaming data. A comparative analysis of classical GNN and dynamic GNN (DGNN — Dynamic Graph Neural Network) modeling methods is conducted, with emphasis on their applicability in trend forecasting, community detection, and anomaly identification. Experiments on publicly available Twitter datasets show that dynamic GNNs achieve 8–10% higher accuracy compared to their static counterparts. Particular attention is paid to issues of scalability, privacy, and model interpretability. The research findings can be applied to develop secure and adaptive algorithms for real-time analysis of social networks.

Keywords: dynamic graph neural networks, temporal graphs, GNN, evolution modeling, real-time data processing, scalability, interpretability, DGNN, trend forecasting, community detection.

Аннотация. Социальные сети представляют собой сложные динамические системы, где связи между пользователями и сущностями (посты, комментарии, сообщества) изменяются в реальном времени. Моделирование подобных временных структур, с целью дальнейшего анализа, с помощью классических методов математического моделирования затруднительно ввиду высокой размерности и необходимости учета аддитивности данных. Традиционные же графовые нейронные сети (GNN — Graph Neural Network) демонстрируют высокую эффективность в анализе статических графов [1–3], однако их применение ограничено при работе с временными данными, требующими адаптации к эволюции структуры графа. В статье рассматриваются современные подходы к разработке динамических GNN, способных моделировать изменения в социальных сетях с учетом временных паттернов. Предложена архитектура, сочетающая рекуррентные механизмы и агрегацию соседей для обработки потоковых данных. Проведен сравнительный анализ методов классического GNN моделирования и динамического GNN (DGNN — Dynamic Graph Neural Network) моделирования, с акцентом на их применимость в задачах прогнозирования трендов, выявления сообществ и обнаружения аномалий. Эксперименты на реальных открытых данных Twitter, показали, что динамические GNN обеспечивают на 5–10 % более высокую точность по сравнению с статическими аналогами. Особое внимание уделено проблемам масштабируемости, приватности и интерпретируемости моделей. Результаты исследования могут быть использованы для создания безопасных и адаптивных алгоритмов анализа социальных сетей в реальном времени.

Ключевые слова: динамические графовые нейронные сети, временные графы, GNN, моделирование эволюции, обработка данных в реальном времени, масштабируемость, интерпретируемость, DGNN, прогнозирование трендов, выявление сообществ.

Введение

Социальные сети стали неотъемлемой частью цифровой экосистемы, генерируя огромные объемы данных с временной динамикой. Ежедневно тысячи пользователей используют социальные сети в различных аспектах, от простого общения друг с другом, до создания новостных групп или групп по интересам, работы и учебы. Традиционные методы анализа подобных динамических систем, такие как статистические модели, математические модели основанные на численных методах оптимизации сложных систем дифференциальных уравнений и дифференциальных уравнений в частных производных, не справляются с подобными задачами ввиду высокой размерности данных, нелинейности связей и высокой динамики изменения структуры такой системы. Для решения подобной области задач существуют классические GNN (Graph Neural Network), которые не учитывают эволюцию связей, вследствие чего снижается их эффективность.

Для решения подобных задач существует возможность использовать DGNN (Dynamic Graph Neural Network), которая позволяет моделировать изменения в графах, сохраняя контекстуальную информацию о прошлых состояниях.

Цель данной работы — разработать архитектуру DGNN, способную адаптироваться к изменениям структуры социальных сетей в реальном времени, и исследовать ее применимость в задачах анализа данных.

Обзор существующих методов

1. Статические GNN

К классу статических GNN подходов анализа можно отнести такие архитектуры сетей, как сверточные графовые нейронные сети, или GCN (Graph Convolutional Network) и графовые нейронные сети с механизмом самовнимания, или GAT (Graph Attention Network). Такие нейронные сети предназначены для анализа графовых данных, имеющих постоянные связи и не изменяемых с течением времени. В связи с этим они не подходят для обработки данных в реальном времени, ввиду изменения структуры связей данных, что может привести к потере важных признаков, или поиску связей там, где они отсутствуют, что может повлечь за собой неправильный результат или поспособствовать переобучению модели.

2. Динамические GNN

Для решения проблемы изменения связей между объектами необходимо производить обучение моделей на временных состояниях связей и узлов графа [4–7]. Подобный механизм применяется в структурах связанных

с динамическими GNN. Примерами таких архитектур являются:

- DYN-GNN — используют рекуррентные механизмы, например Long Short Term Memory (LSTM) для обработки временных графов. Таким образом они учитывают предыдущие состояния графа для определения следующего состояния;
- TGAT (Temporal Graph Attention Network) — вид архитектуры, где реализован механизм внимания на временные связи между узлами. Таким образом предсказание состояния каждого узла строится с учетом временных связей, их появлениями и исчезновениями.

3. Проблемы существующих подходов

При имеющихся достоинствах, такие нейронные сети имеют свой ряд недостатков:

- Масштабируемость — чем больше граф и чем больше в нем связей, тем сложнее обрабатывать связи в данных. Также чем больше граф, тем больше в нем связей, которые необходимо учитывать при обучении и построении прогноза. Но не все важные связи могут быть учтены.
- Ограниченная интерпретируемость решений характеризуется тем, что обработка графа является сложной задачей, и нейронная сеть не предоставляет механизмов и сэмплов данных, на основе которых был сделан прогноз.

Методология

1. Формальная модель динамического графа

Граф $G(t) = (V(t), E(t))$, где $V(t)$ и $E(t)$ — узлы и ребра в момент времени t . Динамика такой системы описывается, как следующая последовательность событий $\{e_1, e_2, \dots, e_n\}$, в которой каждое событие e_i изменяет структуру графа путем добавления или удаления ребер между узлами, или самих узлов.

2. Архитектура динамической GNN

На основании данных архитектура модели должна состоять из нескольких компонент, которые способствуют правильности предсказываемого значения. Модуль временной агрегации необходим для кодирования историчности узлов и связей, чтобы учитывать изменения не только на текущем и предыдущем временном состоянии, но и на протяжении всей жизни графа. Обычно для решения такой задачи используется решение, основанное на LSTM [8].

Механизм внимания позволяет управлять весами модели, которые могут интерпретироваться как степень влияния связанных узлов друг на друга. Процедура взве-

шивания соседей позволяет управлять силами их взаимодействия друг с другом на основе временной близости, то есть чем ближе сосед по времени к текущему узлу, тем больше влияния он имеет на предсказываемое состояние рассматриваемого узла.

Для повышения масштабирования на больших графах и обработки долгих цепочек событий предусматривается механизм обновления графа. Такой подход позволяет обновлять структуру графа относительно временной шкалы, без необходимости полного его перестроения и значительно экономит ресурсы в процессе обучения.

Математически это имеет следующее обоснование:

$$h_v^{(t)} = \sigma \left(W^{(t)} \cdot AGG \left(\left\{ h_v^{(t-\Delta t)} : u \in N(v) \right\} \cup \left\{ h_v^{(t-\Delta t)} \right\} \right) \right)$$

где Δt — смещение по времени относительно предсказываемого момента времени, σ — функция активации нейронов, $W^{(t)}$ — матрица весов, отражающая влияние соседей на состояние предсказываемого узла, AGG — некоторая агрегирующая функция, например сумма, среднее или максимальное из подмножества соседей и предыдущего состояния узла.

Таким образом предсказываемое состояние узла учитывает временные связи, их наличие или отсутствие, а также коэффициенты влияния соседей на итоговое значение.

Эксперименты

В качестве экспериментальных данных был взят открытый датасет от Twitter состоящий из 1.6 миллионов твитов разных пользователей. Основной целью является классификация поста в зависимости от его содержимого по стилистической окраске текста на нейтральный, позитивный или негативный.

Сравнительный анализ статической GNN, в частности архитектуры реализующей сверточную графовую нейронную сеть, или GCN (Graph Convolutional Network) и динамическую графовую нейронную сеть T-GAT (Temporal Graph Attention Network) поведился на метриках точности (Accuracy) и F1-меры (F1-Score), учитывающей несбалансированность классов. По итогам тестов получены следующие показатели исследуемых моделей:

Таблица 1.

Результаты сравнительной характеристики

Модель	Accuracy	F-1 Score
GCN	0.82	0.79
T-GAT	0.87	0.84

Предложенная модель динамической графовой нейронной сети превзошла статическую GCN на 5 % по точности и F1-мере.

Такой результат получен ввиду того, что статические модели не учитывают временные зависимости и последовательности. Также стоит отметить, что в случае статических моделей, чем выше размерность графа данных, тем больше ресурсов необходимо на его обработку, из-за чего возникает проблема масштабируемости, которую в свою очередь решают динамические виды графовых нейронных сетей.

Преимущества и недостатки динамических GNN

В качестве преимуществ динамических моделей можно отметить адаптивность к изменениям внутри моделируемой динамической системы, ввиду обработки временных меток и коррекции весов на их основе [9]. Также такие нейронные сети в некотором роде решают проблему масштабируемости, с которой сталкиваются статические виды GNN [10].

Из недостатков таких сетей, можно отметить сильную зависимость на качество временных меток, то есть чем выше гранулярность временных меток в данных, тем более ресурсоемки будут вычисления, в то время как при низкой гранулярности, точность прогнозирования будет значительно ниже. Другим немаловажным ограничением является выбор обучающих сэмплов, или подграфов, на основе которых будет производиться обучение модели, поскольку это такая же ресурсоемкая процедура, при неправильном выборе которой могут быть потеряны важные зависимости между соседями, или выбраны зависимости, которых на текущий момент уже не существует.

Заключение

Таким образом, предложенная архитектура динамических графовых нейронных сетей позволила обработать большой граф данных, при этом учитывая пространственно-временные связи между соседями, что повлияло на повышение точности предсказания, относительно своего статического аналога.

Для решения аналогичной задачи классическими статистическими методами и методами математического моделированием ушло бы значительно больше времени, поскольку подобная динамическая система имеет большую размерность и огромное количество пространственно-временных связей, которые дополнительно изменяются с течением времени.

ЛИТЕРАТУРА

1. Gromozdov D.R. Graph Convolutional Neural Network Model Robustness under New Data Provisioning / D.R. Gromozdov // Science, Engineering and Business: Conference Proceedings of the VI Interacademic conference for graduate students and young researchers, Moscow, 16–18 April 2024. — Moscow: Bauman Moscow State Technical University (National Research University), 2024. — P. 116–123. — EDN MAMBLM.
2. Лунева Е.Е. Графовые нейронные сети в задачах продленной аутентификации групповых профилей социальных сетей / Е. Е. Лунева, П. И. Банокин // System Analysis and Mathematical Modeling. — 2024. — Vol. 6, № 3. — С. 300–311. — DOI 10.17150/2713-1734.2024.6(3).300-311. — EDN LAQQHA.
3. A study of graph neural networks for link prediction on vulnerability to membership attacks / D.S. Shaikhislamov, K. Lukyanov, N.N. Severin [et al.] // Записки научных семинаров Санкт-Петербургского отделения математического института им. В.А. Стеклова РАН. — 2023. — Т. 530. — С. 113–127. — EDN VLFVUM.
4. Козлов М.В. Возможности использования графовых нейронных сетей как интеллектуальных помощников при построении динамических когнитивных карт / М.В. Козлов, А.В. Михеев, М.П. Дьякович // Современные технологии и научно-технический прогресс. — 2024. — № 11. — С. 137–139. — EDN APHCJJ.
5. Лыченко Н.М. Модели прогноза качества воздуха на основе методов машинного обучения с учетом временных и пространственных зависимостей / Н.М. Лыченко // Проблемы автоматизации и управления. — 2024. — № 3(51). — С. 136–145. — EDN PJSEPL.
6. Predicting epidemic threshold in complex networks by graph neural network / Wu Wang, C. Li, Bo Qu, X. Li // Chaos (Woodbury, N.Y.). — 2024. — Vol. 34, No. 6. — DOI 10.1063/5.0209912. — EDN ELLMMZ.
7. Sighencea B.I. D-STGCN: Dynamic Pedestrian Trajectory Prediction Using Spatio-Temporal Graph Convolutional Networks / B.I. Sighencea, I.R. Stanciu, C.D. Căleanu // Electronics. — 2023. — Vol. 12, No. 3. — P. 611. — DOI 10.3390/electronics12030611. — EDN DWYGVF.
8. DGA-GNN: Dynamic Grouping Aggregation GNN for Fraud Detection / M. Duan, T. Zheng, Ya. Gao [et al.] // Proceedings of the AAAI Conference on Artificial Intelligence. — 2024. — Vol. 38, No. 10. — P. 11820–11828. — DOI 10.1609/aaai.v38i10.29067. — EDN DNHXVO.
9. Ван Ц. Дополненная внимательная GNN для извлечения связей / Ц. Ван // МНСК-2022: Материалы 60-й Международной научной студенческой конференции, Новосибирск, 10–20 апреля 2022 года. — Новосибирск: Новосибирский национальный исследовательский государственный университет, 2022. — С. 192. — EDN ZAKDTA.
10. Romanova A. Enhancing NLP through GNN-Driven Knowledge Graph Rewiring and Document Classification / A. Romanova // Conference of Open Innovations Association, FRUCT. — 2024. — No. 35. — P. 579–587. — EDN RBVSSL.

© Попов Валерий Владиславович (popov.v.v.v.v.p@gmail.com)

Журнал «Современная наука: актуальные проблемы теории и практики»