

На правах рукописи

马辉

МА ЧЖАНЬЦЗЮНЬ

**МЕТОДЫ ОБУЧЕНИЯ ГРАФА ЗНАНИЙ НА ОСНОВЕ ОПТИМИЗАЦИИ
СТРУКТУРЫ ГРАФА И ХЭШИРОВАНИЯ**

АВТОРЕФЕРАТ

диссертации на соискание ученой степени кандидата технических наук

2.3.1 - Системный анализ, управление и обработка информации, статистика

Красноярск 2026

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева», г. Красноярск.

Научный руководитель: доктор технических наук, профессор
Казаковцев Лев Александрович

Официальные оппоненты: **Пимонов Александр Григорьевич**,
доктор технических наук, профессор,
ФГБОУ ВО «Кузбасский государственный
технический университет
имени Т. Ф. Горбачева» (КузГТУ),
г. Кемерово, заведующий кафедрой
прикладных информационных технологий

Царев Роман Юрьевич,
кандидат технических наук, доцент,
ФГБОУ ВО «МИРЭА – Российский
технологический университет» (РТУ
МИРЭА), г. Москва, доцент кафедры высшей
математики института искусственного
интеллекта

Ведущая организация: федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный технический университет»

Защита состоится 29 мая 2026 года в 15:00 часов на заседании диссертационного совета 24.2.403.01, созданного на базе ФГБОУ ВО «Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева» по адресу: 660037, г. Красноярск, пр. им. газеты «Красноярский рабочий» 31, зал заседаний диссертационного совета, ауд. Л-205

С диссертацией можно ознакомиться в научной библиотеке ФГБОУ ВО «Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева» и на сайте <https://www.sibsau.ru>.

Отзывы на автореферат в двух экземплярах, заверенные печатью, просим отправлять по адресу: 660037, Россия, г. Красноярск, просп. им. газеты «Красноярский рабочий», 31, Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева» (СибГУ им. М.Ф. Решетнева), Диссертационный совет E-mail: dissovet@sibsau.ru

Автореферат разослан « ___ » _____ 2026 г.

Ученый секретарь
диссертационного совета

Панфилов Илья Александрович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность. Современные системы искусственного интеллекта, такие как поисковые системы, рекомендательные сервисы и комплексы вопрос-ответ, опираются на структурированное представление знаний о реальном мире. Графы знаний, также известные как семантические сети или онтологии, решают эту задачу, описывая связи между сущностями с помощью графовых структур, что позволяет моделировать сложные семантические отношения. Графовые нейронные сети (GNN) эффективно обучают представления сущностей, распространяя информацию по структуре графа через механизмы передачи сообщений. Однако существующие методы имеют ограничения: Во-первых, автоматически создаваемые графы знаний содержат множество несвязанных подграфов и изолированных узлов, что приводит к искаженным представлениям сущностей с «длинным хвостом» и ограничивает возможности извлечения глобальной структурной информации. Во-вторых, современные графы знаний содержат мультимодальные данные, но существующие методы агрегируют только отдельные признаки триплетов, игнорируя взаимодействия сущностей внутри структуры графа. Поэтому разработка алгоритмов обучения представлений, которые балансируют между характеристиками структуры графа и мультимодальным слиянием, имеет важное исследовательское значение. В-третьих, персонализированные рекомендации должны обрабатывать комбинации продуктов (сущностей), а не отдельные продукты, что приводит к экспоненциальному росту пространства комбинаций и требует разработки более эффективных алгоритмов для вычисления комбинаций. Технологические прорывы в области графов знаний могут улучшить интерпретируемость рассуждений в таких областях, как медицинская диагностика; технологические прорывы в области рекомендательных систем могут вырабатывать предложения сопутствующих услуг и товаров в реальном времени для платформ электронной коммерции с сотнями миллионов пользователей. Вместе они способствуют совместной эволюции когнитивного интеллекта и интеллекта принятия решений.

Степень разработанности тематики. В области обучения представлению графов знаний исследователи предложили различные методы и алгоритмы для решения таких проблем, как несвязность и мультимодальное слияние. Среди методов, основанных на трансляции, в методе TransE Бордеса и др. (2013) впервые применена парадигма моделирования векторной трансляции. Метод TransH Ванга и др. (2014) обрабатывает сложные отношения посредством проекции гиперплоскости, TransR Лина и др. (2015) вводит пространство, специфичное для отношений, а RotatE Суна и др. (2019) моделирует отношения как сложные операции вращения в пространстве. Методы, основанные на тензорной декомпозиции, включают RESCAL Никеля и др. (2011), декомпозицию Tucker, упомянутое в работе Колды и др. (2009) и ComplEx Труйона и др. (2016), использующий сложные вложения. Что касается методов нейронных сетей, Деттмерс и др. В работе ConvE (2018) используются сверточные сети, а в работе НурЕР (2019) Балазевича и др. используются гиперсети для генерации сверточных ядер, специфичных для отношений. В области графовых нейронных сетей

Шлихткрулл и др. (2018) впервые реализовали реляционную свертку графов в RGCN, а Вашишт и др. (2020) улучшили производительность за счет совместного встраивания сущностей и отношений. Методы на основе трансформеров, такие как KG-BERT (2019) Яо и др., CoKE (2019) Ванга и др. и MKGformer (2021) Лю и др., используют предварительно обученные языковые модели для мультимодального рассуждения. Однако существующие графы знаний часто неполны из-за проблем с шумом, создаваемых алгоритмами ручного или автоматического построения, что приводит к несвязным структурам графов. Эта проблема препятствует эффективному использованию существующих графовых нейронных сетей, основанных на принципах распространения информации, в представлении графов знаний, то есть они не могут эффективно улавливать полную структурную информацию графов знаний. Современное обучение представлению графов знаний на основе графовых нейронных сетей в основном ориентировано на одномодальные графы знаний. Мультимодальные взаимодействия в графах фрагментированы, и сущности в графах знаний не ограничиваются текстовыми описаниями, но могут также включать информацию в нескольких модальностях, таких как изображения и речь. В области рекомендаций пакетов услуг Сар Шалом и др. (2016) улучшили показатели кликабельности пользователей с помощью коллаборативной фильтрации. Лю Г. и др. (2017) оптимизировали производительность рекомендаций пакетов услуг с помощью своей модели VRM. Вэй П. и др. (2022) улучшили полноту рекомендаций с помощью своего алгоритма персонализированного объединения. Вейт А. и др. (2015) использовали сиамские сети для повышения точности рекомендаций пакетов услуг. Тансенг и др. В работе (2020) значительно улучшен эффект рекомендаций по комплектации с помощью сквозного метода. Однако в реальных системах очень большое количество элементов обычно увеличивает пространство комбинаций, что делает вычислительную эффективность критически важной проблемой. Поэтому обучение несвязанных графов, слияние мультимодальных графов и обеспечение высокой производительности рекомендательных систем остаются актуальными задачами.

Объектом исследования являются задачи обучения графов знаний, **предметом исследования** являются алгоритмы для решения данных задач.

Цель исследования – повышение точности и вычислительной эффективности решения задач дополнения графов знаний и комбинаторных рекомендаций за счет разработки новых методов обучения представлений, учитывающих структурную несвязность, мультимодальность информации и необходимость работы в реальном времени.

Задачи исследования:

1. Разработать адаптивный алгоритм внимания для графа знаний (Dual-FusionKG), основанный на максимизации взаимной информации, улучшающий способность к обучению представлений несвязных графов посредством структурно-семантического двухканального слияния и исследовать его эффективность по показателям MRR и Hits@10 на наборах данных FB15k, WN18, FB15k237 и WN18RR;

2. Разработать алгоритм, основанный на слиянии тензоров низкого ранга (HyperFusion-Net) с использованием сложных пространственных операций

взаимной корреляции для моделирования направленности мультимодальных связей, повышающий эффективность кросс-модального взаимодействия по показателям Hit@10 и MRR на наборах данных FB15k-237, WN18RR, DB15K и YAGO15K;

3. Разработать алгоритм, совмещающий использование взвешенного хэша и алгоритм вероятностного динамического кодирования (MCHM-Net), сокращающий время ответа рекомендательной системы в режиме реального времени с помощью механизма поиска по двоичным векторам, исследовать эффективность на наборе данных Polyvore-U по показателям AUC (площадь под кривой) и NDCG (нормализованный дисконтированный кумулятивный прирост).

Новые научные результаты:

1. Разработан новый алгоритм Dual-FusionKG для обучения представлений в несвязных статических графах знаний, который, в отличие от известных методов, основан на использовании двухканальной архитектуры сети внимания (ERGAT) с максимизацией взаимной информации между локальными и глобальными представлениями графа, что позволяет эффективно извлекать структурную информацию даже при отсутствии прямых связей между сущностями.

2. Предложен новый алгоритм HyperFusion-Net для обучения представлений мультимодальных графов знаний, основанный на низкоранговом тензорном слиянии текстовых, визуальных и числовых признаков с последующей агрегацией структурной информации через сеть внимания ERGAT, что обеспечивает учет как межмодальных взаимодействий, так и направленности связей в графе.

3. Разработан новый алгоритм персонализированных комбинаторных рекомендаций MCHM-Net, использующий взвешенное хэширование и вероятностное кодирование на основе распределения Бернулли для представления пользователей и элементов в виде бинарных векторов, что позволяет достигать микросекундного времени отклика за счет табличного поиска при сохранении высокой точности рекомендаций.

Теоретическая значимость. Результаты обогащают методы структурно-семантического слияния, моделирования сложного пространства, мультимодального взаимодействия и эффективных рекомендаций хэша. В Dual-FusionKG впервые вводится механизм максимальной взаимной информации на графах для несвязных графов знаний, расширяя границы применения методов, основанных на взаимной информации в обучении графовых структур; HyperFusion-Net конструирует реляционную структуру моделирования, основанную на комплексном пространстве, эффективно объединяя технологию разложения тензоров низкого ранга и предоставляет новую теоретическую перспективу для многомодального слияния; а MCHM-Net вводит механизм совместимости взвешенного хэша и вероятностного динамического кодирования в комбинаторных рекомендациях, что решает ключевую проблему недостаточной гибкости традиционных методов хэширования. Вместе эти подходы способствуют развитию теории обучения представлений в области мультимодального слияния информации, моделирования структурной неоднородности и эффективных рассуждений.

Практическая значимость. Предложенные алгоритмы имеют ценность в практических задачах, таких как завершение графа знаний и комбинаторные

рекомендации. Так, алгоритм DualFusion-KG применим для завершения неполной информации в областях здравоохранения, финансов и т. д.; алгоритм HyperFusion-Net, способный точно моделировать семантические различия в мультимодальных средах данных, применим для совместного вывода текстовых, графических и структурных данных в социальных сетях, в интеллектуальных системах вопрос-ответ и в других областях; алгоритм MCHM-Net за счет способности точно моделировать семантические различия, подходит для совместного рассуждения о текстовых, графических и структурных данных в социальных сетях, интеллектуальных системах вопрос-ответ и др., повышая производительность и масштабируемость комбинированной системы рекомендаций за счет эффективного и динамического механизма хеширования, который широко применим в электронной коммерции, коротких видео, новостных рассылках и других сценариях приложений с высокими требованиями к реальному времени и персонализации. Алгоритм MCHM-NET внедрен в эксплуатацию составе рекомендательной системы технологической компанией "Бэйцзин Циху Кэцзи".

Методы исследования: В диссертации применяются методы из следующих научных направлений: теория графов и машинное обучение на графах (в частности, графовые нейронные сети); многомодальное обучение и слияние данных (использование предобученных моделей VGG-16, BERT, а также предложенный метод низкорангового тензорного слияния HyperFusion-Net); приближенный поиск и хеширование (методы обучения хешированию, включая детерминированное и вероятностное на основе распределения Бернулли); анализ и обработка данных (статистические методы, метрики оценки AUC, MRR, Hits@K и др.). Конкретные технологии, реализованные в рамках этих направлений, включают алгоритмы Dual-FusionKG, HyperFusion-Net и MCHM-Net, которые легли в основу экспериментальной части работы.

Положения, выносимые на защиту:

1. Алгоритм Dual-FusionKG позволяет улучшить структурную осведомленность при сохранении возможностей семантического моделирования за счет двухканального механизма взаимной информации. По сравнению с традиционными методами графовых нейросетей (такими как RGCN и SACN), он демонстрирует преимущества в задачах завершения (устранения несвязности) графа знаний (предсказание недостающих триплетов и определение подлинности заданных триплетов) по показателям $MRR=0,362$ и $Hits@10$.

2. Алгоритм HyperFusion-Net демонстрирует преимущества в мультимодальном динамическом слиянии и моделировании направления связей, достигая максимального значения $Hits@10=0,694$ в задаче завершения графа знаний в эталонных наборах данных, что представляет собой улучшение в среднем на 9,6% по сравнению с лучшими известными моделями, при этом значение MRR улучшается на 7,2%.

3. Алгоритм MCHM-Net позволяет достигать времени отклика рекомендательной системы на уровне микросекунд, представляя рекомендуемые элементы и пользователей рекомендательной системы в виде двоичных векторов, сохраняя при этом точность персонализированных рекомендаций. На наборе данных Polyvore-U значения AUC (площадь под ROC-кривой) и NDCG достигают

0,9237 и 0,8674 соответственно, что значительно превосходит базовые алгоритмы (такие как FPITF и Vi-LSTM); при использовании алгоритма быстрой таблицы поиска время вычисления одной рекомендации составило всего 1,66 микросекунды.

Соответствие паспорту специальности. Работа соответствует п.4 (разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта).

Апробация результатов. Результаты обсуждались на международных научно-практических конференциях «Инновационные исследования: опыт, проблемы внедрения результатов и пути решения» (г.Уфа, 2025), Database and Big Data Workshop (Москва, 2024), «Актуальные проблемы авиации и космонавтики» (г.Красноярск, 2021 и 2023 гг.), Прикладная физика и математика (AAPM-IV 2025, г.Бухара, Узбекистан), ВНПК «Цифровое общество: научные инициативы и новые вызовы» (2025, н.Москва). V Всероссийская (национальная) научная конференция «Достижения науки и технологий» (ДНиТ-V-2026, Красноярск).

Публикации. Результаты опубликованы в 16 работах, среди которых 7 статей в журналах, рекомендованных ВАК.

Структура и объем диссертации. Диссертация состоит из введения, 4 глав, заключения и приложения. Диссертационная работа представлена на 188 страницах, содержит 28 таблиц, 28 рисунков и приложение. Список литературных источников содержит 173 наименования.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обосновывается актуальность, формулируются цели и задачи, научная новизна, перечисляются методы исследования, теоретическая и практическая значимость результатов.

В **первой главе** предлагается алгоритм двухканального кодирования Dual-FusionKG (рисунок 1), предназначенного для сбора полной структурной и семантической информации в графе знаний.

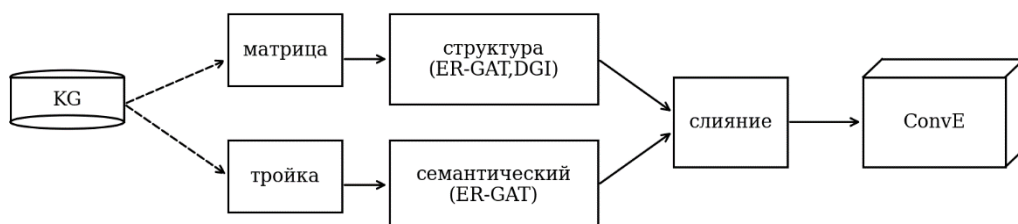


Рисунок 1 – Схема работы алгоритма Dual-FusionKG

Dual-FusionKG – алгоритм кодирования-декодирования для обучения представлению графа знаний, который оперирует двумя независимыми сетями внимания на основе слияния графов сущностей и связей (ER-GAT) в качестве кодера для получения векторов представления (эмбедингов) сущностей и связей и использует ConvE в качестве декодера для прогнозирования оценок триплетов фактов в графе знаний.

При наличии графа знаний входные данные Dual-FusionKG преобразуются в два типа данных для сбора различной информации о графе знаний. Сначала Dual-

FusionKG обрабатывает матрицу смежности графа и собирает полную структурную информацию о знаниях посредством взаимной максимизации информации на первой ER-GAT-сети. Затем на второй ER-GAT-сети семантическая информация извлекается с помощью традиционного обучения с учителем на основе существующих триплетов фактов. После завершения извлечения информации Dual-FusionKG объединяет семантическую и структурную информацию графа знаний для получения окончательного встраивания (эмбединга) сущностей и связей. Наконец, ConvE используется в качестве декодера для получения оценок триплетов фактов, которые необходимо измерить для прогнозирования неполных сущностей, а функция всего алгоритма Dual-FusionKG может быть определена как:

$$f_r(h, t) = f(\text{vec}(f(\text{ER-GAT}[X_h; X_r] * \omega))W)X_t. \quad (1)$$

Здесь $\text{ER-GAT}[X_h; X_r]$ – представление встраивания (эмбединга) с семантической информацией и полной структурной информацией, полученной из ER-GAT, f – операция на полносвязном слое. В контексте ConvE вектор встраивания будет адаптирован как двумерная матрица в качестве входных данных для ConvE-декодера, vec – преобразование (растяжение) карты признаков, сгенерированной обычной свёрткой, в вектор, $*$ – операция свёртки, ω – ядро свёртки в ConvE, а W – матрица весов признаков, извлечённых ConvE.

В триплетах фактов связь может быть сильно коррелирована с хвостовой сущностью или же с головной сущностью. Для учета направленности таких связей и корреляций сущностей предлагается сеть внимания на основе слияния графов сущностей и связей (ER-GAT), которая генерирует промежуточные представления, чувствительные к направлению, путем слияния головной сущности и связи посредством операции взаимной корреляции $\phi(h, r) = h \otimes r$ для каждого триплета (h, r, t) .

ER-GAT использует механизм внимания к связи (рисунок 2) и вводит эмбединг связи в качестве весового коэффициента для определения степени важности связи соседних узлов для целевого узла.

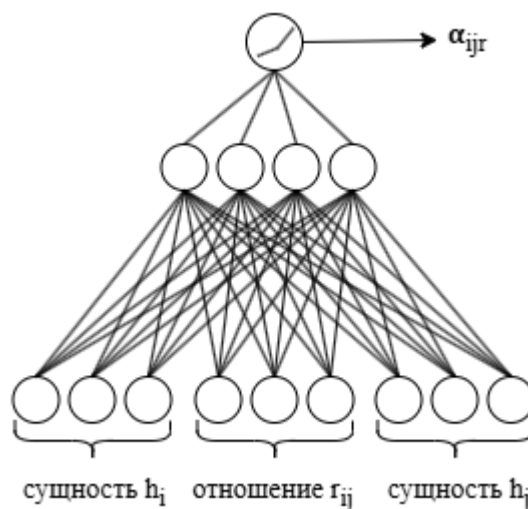


Рисунок 2 – Механизм внимания связи

ER-GAT обновляет скрытую информацию узла с помощью объединенного представления узел-связь, а не просто участвует в функции агрегации через

представление соседних узлов. При этом ER-GAT использует особый механизм внимания связи для определения уровня важности связей соседних узлов по отношению к целевому узлу, обеспечивая адаптивное обновление скрытой информации об узлах. Наконец, ER-GAT создает отдельную матрицу параметров для динамического обновления представления эмбедингов связей, обеспечивая адаптивное обучение представлению отношений в графах с множественными связями.

Функция агрегации ER-GAT математически определяется как:

$$x_i^{(l+1)} = \sigma \left(\sum_{r \in R} \sum_{j \in N} \frac{1}{c_{i,r}} \alpha_{i,j}^{(l)} \varphi(x_i^{(l)}, r_{i,j}^{(l)}) \right). \quad (2)$$

Здесь l представляет l -й уровень ER-GAT, $\alpha_{i,j}^{(l)}$ – оценка уровня внимания узла x_i и связи $r_{i,j}$ с соседним узлом x_j . $c_{i,r} = |N_i^r|$ – коэффициент нормализации, который определяет соседние узлы, с которыми узел i соединен связью r (рисунок 3).

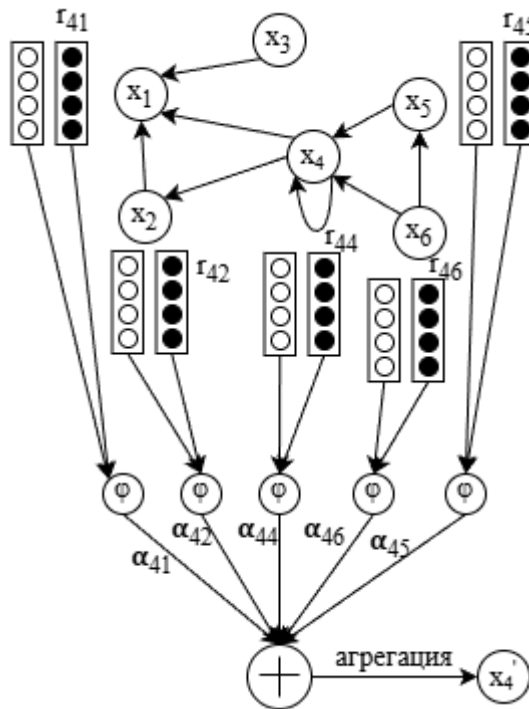


Рисунок 3 – Сеть внимания графа слияния сущностей и связей

Deep Graph Infomax (DGI) – метод обучения без учителя, основанный на представлении графа, который в основном используется для изучения вложенных представлений малой размерности узлов графа или всего графа. Основная идея заключается в получении ключевой информации о структуре и свойствах узлов графа путем максимизации взаимной информации между глобальным представлением графа и локальным представлением узлов.

Цель DGI – максимизировать взаимную информацию между глобальным представлением графа и локальным представлением узлов. Взаимная информация измеряет зависимость между ними, а максимизация взаимной информации означает, что глобальное представление должно содержать как можно больше критически важной информации о локальных узлах. DGI обучает алгоритм различать эти два типа, создавая положительные примеры (исходный граф) и примеры отрицательные (искаженный граф, например, случайно нарушенные

свойства узлов или структура ребер). Глобальные и локальные представления положительных примеров должны быть схожими, а представления отрицательных образцов должны существенно различаться.

На рисунке 4 представлено графическое изображение оценки максимальной взаимной информации с помощью ER-GAT. Здесь $X = (\bar{x}_n, \bar{x}_r)$ – начальный случайный эмбединг графа знаний, где \bar{x}_n и \bar{x}_r – начальные эмбединги сущностей и связей, $Z = (\bar{z}_n, \bar{z}_r)$ – представление эмбедингов, изученных ER-GAT на основе X , где \bar{z}_n и \bar{z}_r распределения обозначают изученные эмбединги сущностей и отношений, $\tilde{X} = (\tilde{x}_n, \tilde{x}_r)$ – начальный случайный эмбединг отрицательных образцов, а $\tilde{Z} = (\tilde{z}_n, \tilde{z}_r)$ – представление эмбедингов, изученных ER-GAT, на основе \tilde{X} .

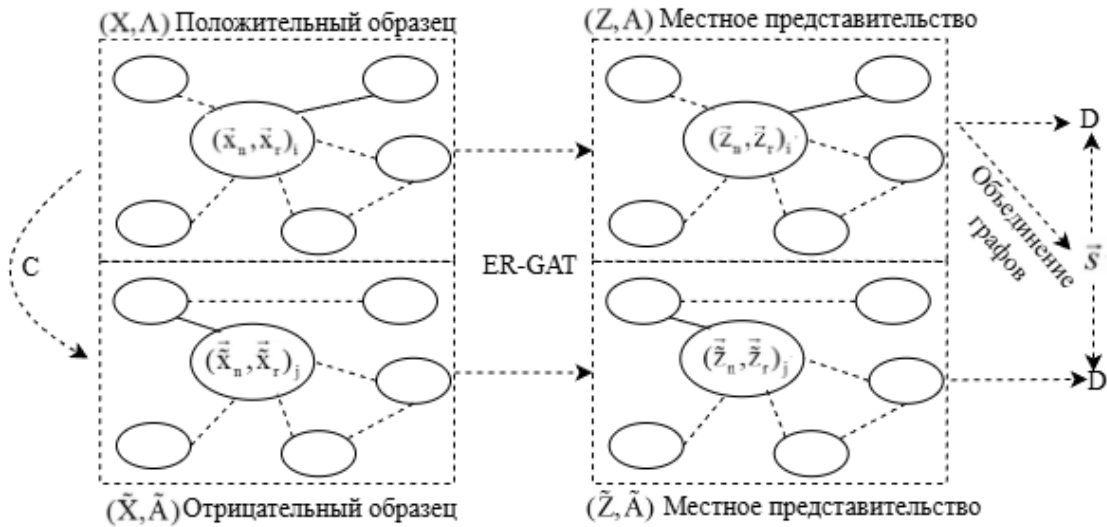


Рисунок 4 – Максимальная взаимная информация на графах на основе ER-GAT

Благодаря операции слияния сущностей и связей и механизму внимания связи в ER-GAT каждая сущность в графе может получить информацию об окружающих ее соседях через функцию агрегации, поэтому Z можно рассматривать как локальное представление графа знаний. Для получения структурной информации на глобальном уровне Dual-FusionKG суммирует полученное локальное представление Z в глобальное представление, используя функцию объединения графов R :

$$\tilde{S} = \frac{1}{|N|} \sum_i^{|N|} (z_n)_i. \quad (3)$$

Здесь $|N|$ – количество сущностей, а $(z_n)_i$ – эмбединг i -й сущности, полученное на основе представления ER-GAT для узла n . Агрегируя всю локальную информацию, можно эффективно получить глобальную информацию графа знаний.

Dual-FusionKG использует дискриминатор D , где выполняется оценка отрицательной выборки. Представление Z и связанная с ним информация рассматриваются как положительная локально-глобальная структурная информация, что подразумевает выборку из совместного распределения локальных и глобальных представлений. Затем отрицательные выборки получают из исходного графа знаний с помощью механизма прерывания C :

$$(\tilde{X}, \tilde{A}) = C(X, A). \quad (4)$$

Отрицательные примеры с правильным значением будут рассматриваться как отрицательная локально-глобальная структурная информация, т.е. они выбираются из произведения распределений рёбер локального и глобального представлений. Граф с отрицательной выборкой имеет точно такой же состав узлов, как и исходный граф, но позиции узлов меняются по сравнению с исходным графом, что приводит к отличающемуся локальному представлению. Когда оценки Dual-FusionKG максимизируют взаимную информацию локальных и глобальных структурных представлений для объединения графов, узлы с различными компонентами связности будут обмениваться информацией для лучшего представления. Подводя итог, Dual-FusionKG получает локальное представление Z исходного графа с помощью прохода (X, A) через сеть ER-GAT, а локальное представление \tilde{Z} графа с отрицательной выборкой с помощью прохода (\tilde{X}, \tilde{A}) через ER-GAT. Следуя идее оценки максимальной взаимной информации, дискриминатор D задается как билинейная функция:

$$D(\vec{z}_i, \vec{s}) = \sigma(\vec{z}_i \vec{W} \vec{s}) \quad (5)$$

Здесь $\vec{z}_i = (\vec{z}_n, \vec{z}_r)_i$. Максимизация оценки взаимной информации на графических знаниях – это процесс слияния локального представления Z в глобальное представление \vec{S} . Полная структурная информация может быть получена как с локальной, так и с глобальной точки зрения, а потери окончательных функций для максимизации оценки взаимной информации на графических знаниях выглядят следующим образом:

$$\begin{aligned} L_{MI} = & \frac{1}{N+M} \left(\sum_{i=1}^N E_{(X,A)} [\log(D((\vec{z}_n, \vec{z}_r)_i, \vec{s}))] \right. \\ & \left. + \left(\sum_{j=1}^M E_{(\tilde{X}, \tilde{A})} [\log(1 - D((\vec{z}_n, \vec{z}_r)_j, \vec{s}))] \right) \right). \quad (6) \end{aligned}$$

Здесь N – число положительных примеров, M – число отрицательных примеров, как правило, $N=M$.

Сравнительная оценка предлагаемого алгоритма Dual-FusionKG производилась на четырёх больших наборах данных для задачи дополнения графа знаний: FB15K, FB15k-237, WN18 и WN18RR (таблица 1).

Таблица 1 – Информация об экспериментальном наборе данных

Название набора данных	FB15K	FB15k-237	WN18	WN18RR
Число сущностей	14981	14541	40943	40943
Число связей	1345	237	18	11
Число триплетов фактов (обучающая выборка)	483142	272115	141442	86835
Число триплетов фактов (валидационная выборка)	50000	17535	5000	3034
Число триплетов фактов (тестовая выборка)	59071	20466	5000	3134
Количество связанных компонентов	9	41	13	430

Сравнение предлагаемого алгоритма Dual-FusionKG проведено с использованием семи известных алгоритмов, основанных на различных декодерах. TransE, RotatE и TorusE – алгоритмы представления знаний, основанные на моделях трансляции, ConvE и HupER – алгоритмы представления знаний,

основанные на сверточных нейронных сетях, а RGCN и SACN основаны на графовых нейронных сетях.

Для экспериментальной оценки (таблицы 2 и 3) были использованы классические метрики для задач дополнения графа знаний: Hits@1, Hits@3 и Hits@10: доля правильных сущностей, попавших в топы 1, 3 и 10 среди всех сущностей-кандидатов.

$$\text{Hits@k} = \sum_i \frac{\Pi(\text{rank}_i \leq k)}{Q}. \quad (7)$$

Здесь $\Pi(*)$ – индикаторная (характеристическая) функция, Q – количество всех тестовых выборок.

Средний рейтинг (MR): средняя доля правильных сущностей, ранжированных среди всех сущностей-кандидатов.

$$\text{MR} = \frac{1}{Q} \sum_i \text{rank}_i. \quad (8)$$

Средний обратный рейтинг (MRR): средняя сумма обратных рейтингов всех сущностей-кандидатов.

$$\text{MRR} = \frac{1}{Q} \sum_i \frac{1}{\text{rank}_i}. \quad (9)$$

Таблица 2 – Сравнительные результаты на наборах данных FB15k-237 и WN18RR

Алгоритм	FB15K237					WN18RR				
	Hits					Hits				
	@10	@3	@1	MR	MRR	@10	@3	@1	MR	MRR
TransE	0.420	0.284	0.174	-	0.257	0.444	0.295	0.027	-	0.182
RotatE	0.533	0.375	0.241	177	0.338	0.571	0.492	0.428	3340	0.476
TorusE	0.485	0.337	0.219	-	0.307	0.534	0.480	0.429	-	0.464
ConvE	0.501	0.356	0.237	244	0.325	0.520	0.440	0.400	4187	0.430
HypER	0.520	0.376	0.252	250	0.341	0.522	0.477	0.436	5798	0.465
R-GCN	0.420	0.230	0.150	-	0.250	-	-	-	-	-
SACN	0.540	0.390	0.260	-	0.350	0.540	0.480	0.430	-	0.470
ERGAT	0.501	0.356	0.237	199	0.325	0.522	0.465	0.421	5131	0.454
Dual-FusionKG	0.549	0.399	0.270	131	0.362	0.543	0.496	0.445	3223	0.491

Таблица 3 – Сравнительные результаты на наборах данных FB15 и WN18

Алгоритм	FB15K					WN18				
	Hits					Hits				
	@10	@3	@1	MR	MRR	@10	@3	@1	MR	MRR
TransE	0.749	0.578	0.297	-	0.463	0.943	0.888	0.113	-	0.495
RotatE	0.884	0.830	0.746	40	0.797	0.955	0.952	0.944	309	0.949
TorusE	0.832	0.771	0.674	-	0.733	0.954	0.950	0.943	-	0.947
ConvE	0.831	0.723	0.558	51	0.657	0.956	0.946	0.935	374	0.943
HypER	0.885	0.829	0.734	44	0.790	0.958	0.955	0.947	431	0.951
R-GCN	0.825	0.736	0.541	-	0.651	0.955	0.928	0.686	-	0.814
ERGAT	0.885	0.822	0.715	71	0.776	0.956	0.929	0.891	478	0.898
Dual-FusionKG	0.896	0.854	0.746	38	0.798	0.957	0.950	0.943	436	0.948

Алгоритм DualFusionKG показывает преимущество лучше на наборах данных (FB15k-237, WN18RR), при этом MRR у DualFusionKG (0,362) на 3,4% выше, чем у SACN (0,350) на FB15k-237, а на WN18RR MRR (0,491) выше, чем у SACN

(0,470) на 4,5%. Причина в том, что DualFusionKG максимизирует сбор глобальной структурной информации посредством взаимной информации, что решает проблему изоляции информации традиционных графовых нейронных сетей (например, SACN) из-за несвязанных компонентов. Два узла не связаны напрямую, но улучшают способность вывода за счет глобального семантического обмена. Алгоритм DualFusionKG имеет меньшее преимущество на наборах данных (FB15K, WN18), но все же лидирует на фоне известных алгоритмов. На FB15K значение MRR у DualFusionKG (0,798) немного выше, чем у RotatE (0,797), а значение Hits@10 (0,896) значительно выше, чем у HyperER (0,885). На WN18 значение Hits@10 у DualFusionKG немного ниже, чем у HyperER (0,957), но MRR близок к оптимальному (0,948). Это может быть связано с тем, что набор данных содержит большое количество обратных связей, которые могут быть эффективно предсказаны традиционными алгоритмами (например, RotatE) с помощью простых закономерностей (например, симметрии), и преимущество сложной структуры DualFusionKG частично нивелируется. В совокупности эти результаты позволяют сделать вывод об эффективности алгоритма DualFusionKG и его более высокой производительности при решении задачи дополнения графа знаний.

Во **второй главе** предлагается алгоритм HyperFusion-Net, который позволяет эффективно собирать мультимодальную информацию и информацию о структуре графа в мультимодальных графах знаний и реализовывать сквозной процесс обучения. HyperFusion-Net реализует обучение представления мультимодальных графов знаний посредством трёх основных модулей: модуля слияния информации, модуля агрегации информации и модуля проверки прогнозов.

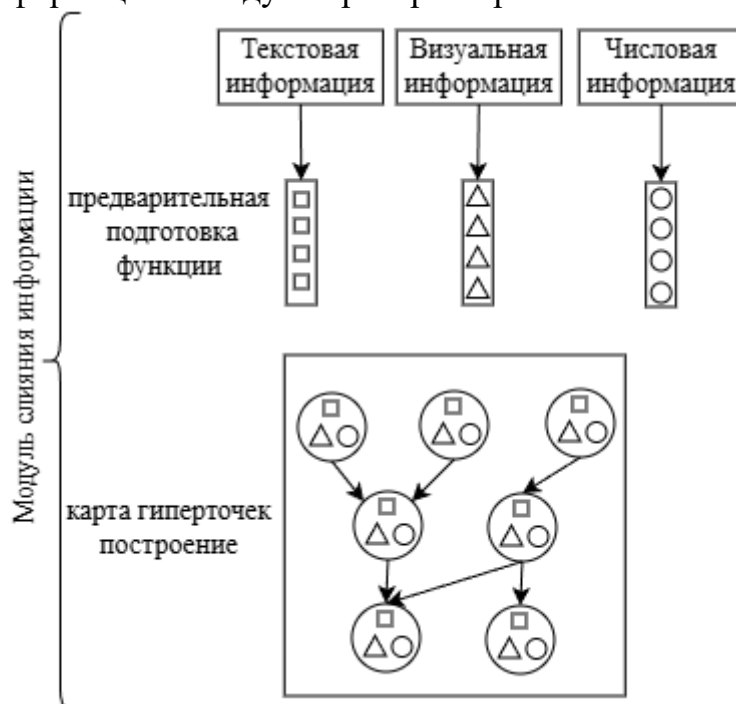


Рисунок 5 – Структура модуля слияния информации отдельных модальностей

Модуль слияния информации (рисунок 5): модуль сначала получает эмбединг текстовой информации с помощью BERT, эмбединг визуальной информации с помощью VGG16 и эмбединг числовой информации с помощью BERT+FC. VGG-16 и BERT, используемые в модуле слияния информации, являются предобученными алгоритмами. После завершения извлечения признаков

предобученных данных модуль слияния информации проецирует все предобученные эмбединги в эмбединги той же размерности. Затем исходный мультимодальный граф знаний преобразуется в гиперточечный граф путем выполнения низкорангового тензорного слияния предобученных эмбедингов отдельных модальностей.

Низкоранговое тензорное слияние – эффективный метод слияния мультимодальной информации, который снижает вычислительную сложность и извлекает ключевые признаки с помощью метода тензорного разложения. Низкоранговое тензорное слияние e для информации трех модальностей можно выразить как

$$e = \left(\sum_{i=1}^k w_t^{(i)} \cdot e_t \right) \circ \left(\sum_{i=1}^k w_v^{(i)} \cdot e_v \right) \circ \left(\sum_{i=1}^k w_n^{(i)} \cdot e_n \right). \quad (10)$$

Здесь « \circ » – поэлементное произведение.

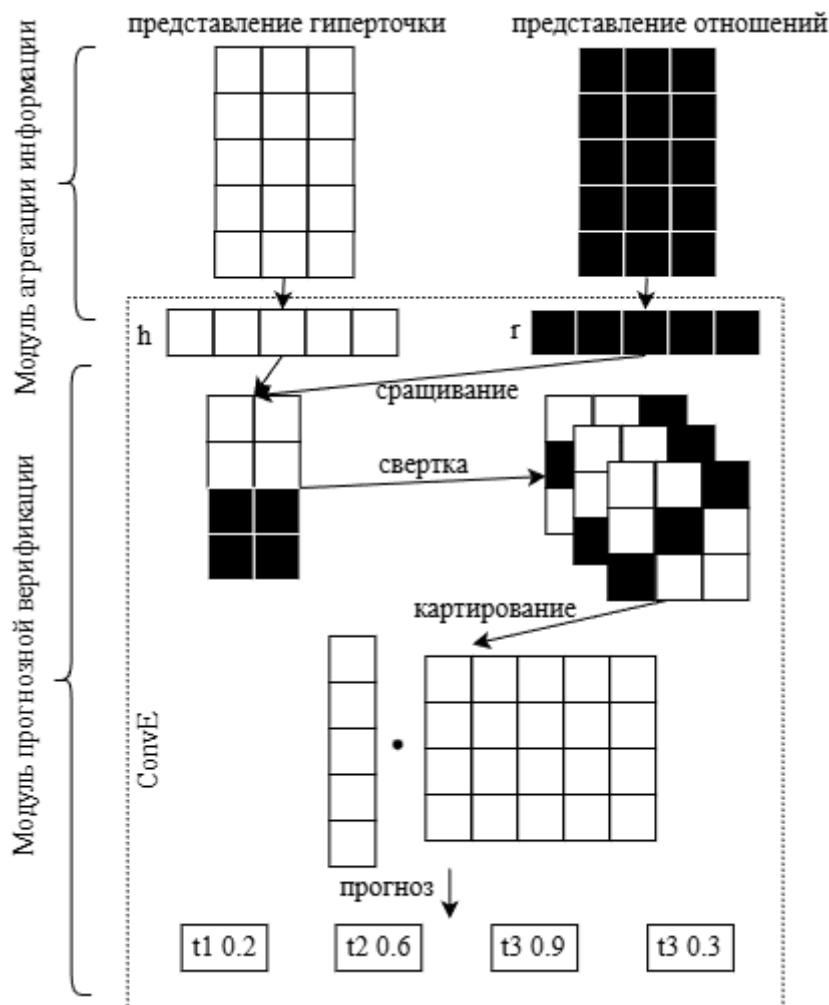


Рисунок 6 – Иллюстрация работы модуля агрегации информации и прогнозирования

В модуле агрегации (рисунок 6) информация о структуре графа собирается путем агрегации информации о гиперточках соседей. Сначала для каждой гиперточки измеряется важность ее соседей с помощью сети внимания на основе слияния сущностей и связей (ER-GAT), предложенной в первой главе. Затем информация о структуре графа обновляется путем слияния информации о сущности с информацией о связях. Затем для каждой пары «сущность-связь»

соседа информация о сущности и информация о связях объединяются ER-GAT для обновления представления суперузла. Это позволяет модулю агрегации информации эффективно собирать информацию о структуре графа в мультимодальном графе знаний, одновременно улучшая представление сущностей и связей.

В модуле валидации прогноза в качестве декодера выбран ConvE, а функция оценки для каждой тройки фактов имеет вид:

$$f(\mathbf{h}, \mathbf{r}, \mathbf{t}) = \sigma(\text{vec}([\mathbf{e}_h; \mathbf{r}] * \mathbf{W}) \mathbf{e}_t). \quad (11)$$

Здесь \mathbf{e}_h и \mathbf{e}_t – представления эмбедингов сущностей «головой» и «хвоста» связи соответственно, а \mathbf{r} – реляционное представление (представление связи) эмбедингов между ними, $*$ и ω – операция свёртки и ядро свёртки соответственно, а vec – операцию реконструкции (растяжения) карты признаков в векторы, \mathbf{W} – матрица отображения, а σ – сигмоидальная функция активации. При использовании описанного выше подхода алгоритм может вычислить, является ли факт, составленный между сущностями на основе некоторой связи, истинным или ложным.

В данной работе в качестве функции потерь выбрана бинарная перекрестная энтропия:

$$L(f, t) = -\frac{1}{|F|} \sum_i (t_i \cdot \log(f_i) + (1 - t_i) \cdot \log(1 - f_i)). \quad (12)$$

Здесь $|F|$ – общее количество обученных триплетов фактов, t_i – метка триплета фактов, f_i – оценка прогноза i -й тройки фактов на основе сети внимания мультимодального графа знаний. Значение t равно 1, когда триплет фактов истинен, и 0, когда он ложен.

Сравнение алгоритма HyperFusion-Net проводилось с девятью другими алгоритмами на тестовых наборах данных FB15k-237, WN18RR, DB15K и YAGO15K (таблица 4). Метрики оценки те же, что и в первой главе.

Таблица 4 - Информация о наборах данных (сравнение HyperFusion-Net)

Набор данных	FB15k237	WB18RR	DB15K	YAGO15K
Число сущностей	14541	40943	12842	15404
Число связей	237	11	279	32
Число триплетов фактов	310116	93012	89197	122886
Число сущностей текстовой информации	14541	40943	12842	11199
Число сущностей визуальной информации	13444	-	12837	11194
Число сущностей числовой информации	29395	-	48080	23532
Число атрибутов	116	-	225	7

FB15k-237, DB15K и YAGO15K содержат информацию всех модальностей, а WN18RR содержит только текстовую информацию и информацию о базовых триплетах фактов. Получены результаты представлены в таблице 5.

Таблица 5 – Сравнительные результаты на FB15k-237 и WN18RR

Алгоритм	FB15k-237					WN18RR				
	Hits					Hits				
	MR	MRR	@1	@3	@10	MR	MRR	@1	@3	@10
TransE	357	0.257	0.174	0.284	0.420	-	0.182	0.027	0.295	0.444
RotatE	177	0.338	0.241	0.375	0.533	3340	0.476	0.428	0.492	0.571
JointE	177	0.356	0.262	0.393	0.543	4655	0.537	0.438	0.483	0.537
ConvE	244	0.325	0.237	0.356	0.501	4187	0.430	0.400	0.440	0.520
KMAE	235	0.326	0.240	0.358	0.502	4441	0.448	0.415	0.465	0.524
R-GCN	-	0.250	0.150	0.230	0.420	-	-	-	-	-
WGCN	-	0.350	0.260	0.390	0.540	-	0.470	0.430	0.480	0.540
KBLRN	209	0.309	0.219	0.342	0.493	3871	0.471	0.435	0.479	0.538
MMRFAN	256	0.297	0.205	0.326	0.485	-	-	-	-	-
Hyper Fusion-Net	156	0.366	0.271	0.404	0.542	2685	0.491	0.454	0.503	0.567

На FB15k-237 (мультимодальный набор данных) HyperFusion-Net по значению Hits@10=0.542 существенно опережает известные алгоритмы. Значение MRR достигает 0.366, что на 2.8% выше, чем у ближайшего конкурента (JointE: 0.356). Значение MR (средний рейтинг) равно 156, что существенно ниже, чем у известных алгоритмов (напр, для TransE: 357), т.е. новый алгоритм способен точнее находить требуемую сущность. На WN18RR (только текстовые данные) достигаемое HyperFusion-Net значение Hits@10=0.567 близко к лучшему из известных алгоритмов RotatE (0.571), при этом значение MRR=0.491 существенно лучше, чем у RotatE (0.476), что позволяет предположить, что он более стабилен в сложных рассуждениях на основе связей.

Алгоритм HyperFusion-Net достигает показателя Hits@10=0,694 на наборе данных DB15K, что на 4,5% лучше, чем у лучшего из известных алгоритма WGCN (0,664). Показатель Hits@10 на YAGO15K составляет 0,637, что значительно лучше показателей KBLRN (0,578) и R-GCN (0,461), что подтверждает способность нового алгоритма адаптироваться к различным распределениям данных (таблица 6).

Таблица 6 – Сравнительные результаты дополнения графа знаний (наборы данных DB15K, YAGO15K)

Алгоритм	DB15K					YAGO15K				
	Hits					Hits				
	MR	MRR	@1	@3	@10	MR	MRR	@1	@3	@10
TransE	3	0.257	0.174	0.284	0.420	-	0.182	0.027	0.295	0.444
RotatE	1805	0.52	0.5	0.531	0.551	1732	0.356	0.289	0.391	0.479
R-GCN	2101	0.531	0.479	0.58	0.574	1540	0.304	0.258	0.308	0.461
WGCN	1243	0.601	0.569	0.640	0.664	740	0.454	0.358	0.508	0.611
KBLRN	1357	0.578	0.531	0.623	0.657	1182	0.414	0.335	0.492	0.578
Hyper Fusion-Net	917	0.630	0.597	0.647	0.694	474	0.480	0.406	0.524	0.637

Архитектура алгоритма включает часть, отвечающую за извлечение признаков: для пользовательских данных бинарное представление пользователя получается с помощью one-hot кодирования, признаки изображения извлекаются с

использованием CNN, а семантические признаки извлекаются с помощью текстовой модели.

Хэш-алгоритм на основе функции активации: после вывода признаков пользователей и элементов добавляется слой функции активации ScaleTanH для получения бинарного кодирования:

В третьей главе предлагается алгоритм многокатегорийной хеш-сети (Multi-Category Hashing Matching Network, MCHM-Net) с вероятностной интерпретацией на основе хеш-функции (HF). Вычисления ускоряются за счет двоичного кодирования с сохранением возможности персонализированных рекомендаций.

Архитектура алгоритма содержит несколько модулей:

- Модуль извлечения признаков: для пользовательских данных двоичное представление пользователя получается методом прямого кодирования, признаки изображения извлекаются с помощью сверточной нейронной сети (Convolutional Neural Network, CNN), а семантические признаки извлекаются с помощью текстовой модели.

- Модуль кодирования хеш-функций: для различных категорий элементов разрабатываются независимые хеш-функции для генерации двоичного кодирования.

- Алгоритм хеширования на основе функции активации: добавление функции активации ScaleTanH после вывода признаков пользователей и элементов для получения двоичного кодирования. ScaleTanH – вариант функции активации TanH, где параметр β растет с ростом числа итераций t (рисунок 7).

$$\text{ScaleTanH}(x) = \text{TanH}(\beta_t \cdot x), \beta_t > 0 \quad (13)$$

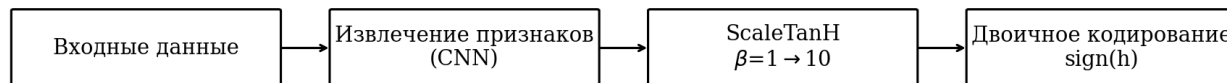


Рисунок 7 – Алгоритм хеширования на основе функции активации

- Алгоритм хеширования, основанный на распределении Бернулли: каждый двоично-кодированный бит рассматривается как независимое распределение Бернулли. Затем генерируется двоичный вектор путём многократной выборки. Для k -го бита вероятность принятия значения 1 равна p_k , вероятность принятия значения -1 равна $1-p_k$. Выводятся параметры вероятности через применение нейронной сети. Каждый бит выбирается независимо для генерации двоичного кода, который увеличивается на n_s , а длина кодирования увеличивается (рис. 8).

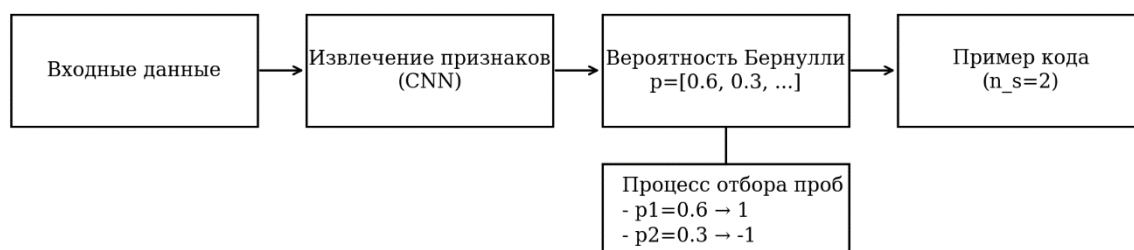


Рисунок 8 – Алгоритм хеширования на основе распределения Бернулли

- Модуль сопоставления: оценка предпочтений пользователя r_{u,O_i} для множества вычисляется с помощью взвешенного внутреннего произведения.

$$r_{u,O_i} = \frac{1}{Z_1} \sum_n h_{i_n,t}^{(n)T} \Lambda_1 h_{u,t}^u + \frac{1}{Z_2} \sum_{n \neq m} h_{i_n,t}^n \Lambda_2 h_{i_m,t}^m \quad (14)$$

Здесь Λ_1 и Λ_2 – неотрицательные диагональные матрицы, используемые для определения важности различных битов в двоичном векторе. В этой функции два члена: первый указывает, насколько пользователь предпочитает множество, а второй описывает, насколько хорошо соответствует само множество.

Целевая функция (функция потерь на основе BPR – байесовского персонализированного рейтинга):

$$L_{\text{rank}} = \sum_{(u,i,j) \in P} \log(1 + \exp(-(r_{u,o_i} - r_{u,o_j}))) \quad (15)$$

Для дальнейшего повышения эффективности строится таблица быстрого поиска путём предварительного вычисления взвешенного расстояния Хэмминга. Пусть длина двоичных векторов равна D . Разделим каждый двоичный вектор на подблоки. Для каждого подблока предварительно вычислим все возможные результаты перестановок и соответствующие им взвешенные расстояния. Для каждого результата k вычислим взвешенное расстояние.

$$v_k = \sum_{i=1}^n w_i \cdot k_i \quad (16)$$

Здесь w_i – вес, а k_i – i -я позиция результата перпозиционного различия (таблица 7).

Таблица 7 – Таблица поиска для взвешенных расстояний Хэмминга

k	v
0	0
1	w_1
2	w_2
3	$w_1 + w_2$
...	...
254	$w_2 + \dots + w_8$
255	$w_1 + \dots + w_8$

- Модуль быстрого поиска: при вычислении взвешенного расстояния Хэмминга два двоичных вектора сначала подвергаются побитовой проверке на несходство для получения результата k , а затем соответствующее взвешенное расстояние получается непосредственно из таблицы поиска. Поскольку сложность поиска по таблице постоянна ($O(1)$), такой подход очень эффективен.

Для двоичного вектора длины D его можно разбить на $s = \lceil D/8 \rceil$ подблоков, и взвешенное расстояние Хэмминга для каждого подблока вычисляется отдельно, а затем окончательное взвешенное расстояние Хэмминга может быть получено путем суммирования расстояний всех подблоков.

Наборы данных с сайта Polyvore (таблица 8) широко используются в исследованиях, связанных с рекомендательными системами комбинаций одежды. В этом исследовании был создан набор данных Polyvore-U, содержащий полные данные о поведении пользователей таких систем. Набор данных сохраняет три основных элемента подбора комплектов одежды – верх, низ и обувь – и разработан с двойной архитектурой данных, основанной на реальных сценариях подбора пар, таких как базовый подбор комплектов (комбинации строго из трех отдельных предметов одежды). Для этого сценария сформирован базовый набор данных, содержащий данные 630 пользователей, разделённый на Polyvore-630 (обучающая

выборка) и Polyvore-53 (тестовая выборка). Сценарию расширенного подбора комплектов (допускающему 1-2 предмета «верх») соответствует расширенный набор данных, содержащий данные 551 пользователя, разделенные на Polyvore-519 (обучающая выборка) и Polyvore-32 (тестовая выборка). В наборе данных используется строгий механизм разделения элементов, гарантирующий отсутствие перекрытия элементов в обучающей и тестовой выборках, а многомерная среда валидации построена на основе стратегии дифференцированной сегментации (существует частичное перекрытие между базовыми и расширенными группами пользователей). В наших экспериментах Polyvore-630 и Polyvore-519 используются в основном для базовой валидации алгоритма, тогда как Polyvore-53 и Polyvore-32 используются в качестве независимых тестовых наборов для оценки способности модели к обобщению на новых пользователей.

Таблица 8 – Информация о наборах данных

	Число комплектов		Число предметов одежды	
	Обучающая выборка	Тестовая выборка	Обучающая выборка	Тестовая выборка
P-630	127326	23054	159729	45505
P-53	10712	1944	20230	4437
P-519	83416	14654	146475	39085
P-32	5133	898	14594	2797

Для сравнительных экспериментов рассматривались следующие известные алгоритмы: SiameseNet, Bi-LSTM, Type-aware, SCE-Net. SiameseNet изучает корреляцию между парами элементов, используя двойную сеть, где сеть изучает признаки, приближая представления схожих элементов друг к другу и отдаляя их друг от друга между разнородными элементами. На основе этой сети оценка набора определяется как среднее значение сходства всех пар элементов в наборе. Bi-LSTM рассматривает набор как последовательность, а выбор элементов как задачу прогнозирования. Совместимость набора оценивается путем вычисления вероятности появления всех элементов. Type-aware разлагает совместимость набора на совместимость между парами элементов. Type-aware разлагает совместимость набора на пары элементов и сопоставляет различные элементы с различными подпространствами для метрического обучения с использованием информации о категориях, в то время как SCE-Net изучает несколько условных эмбедингов элемента для метрического обучения. Вес каждого условного встраивания рассчитывается механизмом внимания.

В ходе эксперимента (таблица 9) были проанализированы две задачи рекомендации. Первая задача – рекомендация одежды, где для каждого пользователя костюмы сортируются в соответствии с его балльной оценкой. Качество сортировки костюмов оценивалось по площади под ROC-кривой (AUC) и нормализованному дисконтированному кумулятивному приросту (NDCG). Вторая задача – задача заполнения пропусков (FITB). В задаче FITB используется неполный комплект предметов одежды, и требуется выбрать один наиболее подходящий предмет из набора кандидатов заданной категории. Также анализируется время выполнения.

AlexNet используется в качестве опорной сети для обучения и извлечения признаков на основе изображений. Для обработки изображений разного

разрешения все полносвязные слои в AlexNet заменяются сверточным слоем, а для получения 4096-мерного вектора признаков добавляется слой усреднения. Для текстовых данных текстовые признаки получаются из описаний и этикеток элементов с помощью существующего алгоритма последовательности, где размерность текстовых признаков составляет 2400.

Для минимизации погрешности, вызванной случайной генерацией отрицательных образцов, каждый алгоритм тестировался 10 раз, и в качестве окончательного результата использовалось среднее значение этих 10 запусков. Каждый из этих тестов регенерировал набор отрицательных образцов, используемых для сравнения.

Таблица 9 – Производительность алгоритма на данных Polyvore-630 и Polyvore-519

Алгоритм	Polyvore-630			Polyvore-519		
	FITB	AUC	NDCG	FITB	AUC	NDCG
SiameseNet	0.5501	0.8123	0.6623	0.5621	0.8315	0.6736
Bi-LSTM	0.5610	0.8211	0.6943	0.5977	0.8569	0.7153
Type-Aware	0.5576	0.8166	0.6756	0.5600	0.8234	0.6801
SCE-Net	0.5858	0.8400	0.7161	0.5763	0.8348	0.6951
MCHM-Net -T(ScaleTanH)	0.5620	0.8686	0.7786	0.5450	0.8666	0.7711
MCHM-Net -T (Deterministic)	0.5718	0.8739	0.7903	0.5655	0.8773	0.7929
MCHM-Net -T ($n_s=3$)	0.5729	0.8747	0.7936	0.5674	0.8793	0.7969
MCHM-Net -V (ScaleTanH)	0.6145	0.8979	0.8178	0.6027	0.8974	0.8106
MCHM-Net -V (Deterministic)	0.6172	0.8992	0.8199	0.6145	0.9032	0.8205
MCHM-Net -V ($n_s=3$)	0.6263	0.9036	0.8294	0.6191	0.9068	0.8272
MCHM-Net-VSE	0.6628	0.9237	0.8674	0.6575	0.9282	0.8722

В таблице 9 «Т» указывает на то, что алгоритм использует только текстовые признаки, «V» указывает на то, что алгоритм использует только признаки изображений, а «VSE» указывает на то, что алгоритм использует информацию обеих модальностей. Сравнение производительности двух предложенных алгоритмов хеширования, алгоритма хеширования на основе функции активации ScaleTanH и алгоритма хеширования на основе распределения Бернулли приведено в алгоритме с использованием одномодальных данных соответственно. Для алгоритма хеширования на основе распределения Бернулли производительность также приводится отдельно для двух подходов бинаризации: первый – детерминированная бинаризация, второй – путем выборки каждого распределения Бернулли. Для мультимодальных данных показана только детерминированная бинаризация на основе распределения Бернулли с лучшей производительностью. ResNet-18 используется в качестве сети извлечения признаков для визуальных данных. По результатам сравнения обнаружено, что предложенный алгоритм хеширования MCHM-Net-VSE превосходит известные алгоритмы по всем использованным метрикам. Новый алгоритм по-прежнему превосходит другие методы, использующие многомодальную информацию, даже при использовании только изображений. Это доказывает эффективность предлагаемого алгоритма MCHM-Net. Он показывает, что, несмотря на повышение вычислительной эффективности за счет метода хеширования, он достигает продвинутого уровня точности рекомендаций; мультимодальная

информация вносит свой вклад в производительность алгоритма. Сравнивая производительность двух модальностей, можно обнаружить, что информация об изображении более полезна в этой задаче, чем текст. Основная причина этого заключается в том, что сопоставление одежды больше ориентировано на визуальное единство, а текстовая информация о каждом предмете ограничена. Эти две причины одновременно позволяют алгоритму, основанному на информации об изображении, достичь более высокой производительности. Производительность алгоритма дополнительно повышается за счет объединения текстовой и графической информации.

Двоичное представление в некоторой степени снижает производительность алгоритмов, но обеспечивает значительное удобство вычислений. Повышение эффективности вычислений двоичных векторов демонстрируется путем сравнения времени, необходимого различным алгоритмам для выполнения расчетов (табл.10).

Таблица 10 – Сравнение времени выполнения алгоритмов

Алгоритм	MCHM-Fast	MCHM-Net	Siamese-Net	Bi-LSTM	Type-Aware	SCE-Net
Время, мкс	1.	27.	38.	1.23	59.1	192
Сравнительное ускорение, раз	-	~16.4	~ 23.4	~741.0	~ 35.6	~115.7

Здесь MCHM-Fast обозначает вычисления с использованием быстрых таблиц поиска, а MCHM-Net – без них. Прежде всего, можно увидеть, что при простой реализации ускорение, основанное на быстрых таблицах поиска, как минимум в 16 раз выше по сравнению с производительностью без использования быстрого алгоритма. Технология хеширования значительно повышает эффективность комбинаторных рекомендаций благодаря двоичному кодированию, что очень полезно для рекомендаций в реальном времени.

ЗАКЛЮЧЕНИЕ

Предложен ряд новых алгоритмов, направленных на повышение эффективности решения задач по дополнению графа знаний за счет снижения вычислительной сложности в мультимодальных комбинаторных пространствах за счет решения проблемы несвязности структуры графа знаний и многомодальности при обучении представлению графа знаний на основе графовых нейронных сетей. Также предложены алгоритмы для реализации системы онлайн-рекомендаций по сочетанию предметов одежды на основе графов знаний. В ходе исследования решены следующие задачи.

1. Предложен алгоритм обучения представлению графа знаний Dual-FusionKG, который объединяет структурную и семантическую информацию. Алгоритм моделирует локальную структурную информацию и троичную семантическую информацию по отдельности, используя двухканальную адаптивную сеть внимания графа (ERGAT), и максимизирует локальное и глобальное представления за счет максимизации взаимной информации. Это эффективно устраняет проблему, связанную с трудностями распространения межсущностной информации в несвязных графах. Экспериментальные результаты показывают, что Dual-FusionKG значительно превосходит существующие алгоритмы на нескольких

эталонных наборах данных, что подтверждает эффективность предлагаемого алгоритма в задаче завершения графа знаний.

2. Предложен алгоритм HyperFusion-Net для динамического моделирования взаимодействий с учетом направленности гетерогенных связей мультимодальных графов знаний посредством низкорангового тензорного слияния со сложными пространственно взаимосвязанными сетями слияния сущностей и связей, что значительно повышает производительность задач дополнения графов знаний. Эксперименты показывают, что алгоритм HyperFusion-Net улучшает показатель Hits@10 до 9,6% на четырех тестовых наборах данных, будучи и устойчив к пропускам модальности, обеспечивая эффективное решение для сложных задач анализа знаний.

3. Предложен алгоритм MCHM-Net, который значительно повышает эффективность и точность комбинаторной рекомендательной системы благодаря методу хеширования и достигает $AUC=0,9237$ при времени отклика на уровне микросекунд, используя алгоритм взвешенного хеширования и вероятностного кодирования Бернулли.

Результаты исследования, представленные в данной работе, открывают несколько направлений для дальнейших исследований. Во-первых, изучение применения методов дистилляции знаний для сжатия модели Dual-FusionKG с целью уменьшения ее пространственной сложности без существенной потери точности является практически важной темой. Во-вторых, с точки зрения расширения мультимодальных методов (HyperFusion-Net), включение дополнительных модальностей (таких как аудиоданные или временные ряды) и изучение более сложных механизмов внимания для учета семантической направленности отношений демонстрируют многообещающие перспективы. Наконец, для алгоритма MCHM-Net ценным направлением является разработка методов «динамического» расширения длины кодирования для адаптации к изменениям пользовательских предпочтений и достижения более высокого качества рекомендаций; в то же время, применение разработанного метода хеширования к другим типам задач рекомендаций, помимо подбора одежды, также имеет широкие перспективы применения.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в журналах, рекомендованных ВАК:

1. Ma, Z. Multimodal knowledge graph attention networks/ Z. Ma, L.A. Kazakovtsev // Системы управления и информационные технологии, №2(100), 2025. С. 58-65 (Перечень ВАК К2).

2. Ma, Z. Online ordering recommendation method based on bipartite graph projection and node2vec algorithm / Z. Ma, M. Gao, L.A.Kazakovtsev // Системы управления и информационные технологии. 2025. № 1 (99). С. 22-27 (Перечень ВАК К2).

3. Ma, Z. The design of clothing combination recommendation system based on hash technology / Z. Ma, M. Gao, Z. Guo // Системы управления и информационные технологии. 2025. № 1 (99). С. 87-94 (Перечень ВАК К2).

4. Guo, Z. Analysis and prediction of data on the basic situation of forest fires / Ma Z, Guo Z, Sopov E.A. // Перспективы науки. 2024. № 5 (176). С. 66-70 (Перечень

ВАК К3).

5. Гао, М. Алгоритм группирования текстовых данных с рекомбинацией по принципу химической реакции / Ч.Ма, М.Гао, М.И.Цепкова, Н.Л.Резова, А.А.Ступина // Системы управления и информационные технологии. 2022. № 3 (89). С. 61-65 (Перечень ВАК К2).

6. Гао, М. Построение динамической гибридной модели рекомендаций и исследования пользователей / М. Гао, Ч. Ма, Л.А.Казаковцев // Сибирский аэрокосмический журнал. 2025. Т. 26, № 4. С. 466–477 (Перечень ВАК К2).

7. Ма, Z. Geomrtry-Aware Graph Neural Network with Vox Embeddings / Ч. Ма, Л. А. Казаковцев // Системы управления и информационные технологии. 2025. №4(102). С. 4-10 (Перечень ВАК К2).

Зарегистрирована программа для ЭВМ

1. Ма, Ч. NCHM-NET. Свидетельство о государственной регистрации программы для ЭВМ / Ч. Ма., Л. А. Казаковцев // Роспатент. рег. № 2026615791 от 03.03.2026, заявл. 23.02.2026.

Прочие публикации:

1. Ма, Z. Design and implementation of tourism visualization attraction recommendation system based on collaborative filtering algorithm / Z. Ma, M. Gao // Innovation Science. – 2025. – No. 2-2. – P. 48-53.

2. Ма, Z. Design and implementation of news recommender system based on collaborative filtering and machine learning / Z. Ma, M. Gao // Innovation Science. – 2025. – No. 2-2. – P. 53-58.

3. Ма, Z. Design and implementation of author - based collaborative filtering book recommendation algorithm / Z. Ma, M. Gao // Инновационные исследования: опыт, проблемы внедрения результатов и пути решения: сборник статей Международной научно-практической конференции, 17 февраля 2025 года, г.Уфа, 2025. – P. 31-38.

4. Ма, Z. A study on weighted collaborative filtering recommendation algorithm based on multi-cluster hybridization / Z. Ma, M. Gao // Innovation Science. – 2025. – No. 2-2. – P. 58-63.

5. Ма, Z. Collaborative filtering-based recommendation strategy for reducing false ratings / Z. Ma, M. Gao, Z. Guo // Innovation Science. – 2024. – No. 1-2. – P. 56-59.

6. Ма, Z. Self-optimizing weighted Bayesian classifier based on heuristic algorithm / Z. Ma, M. Gao // Innovation Science. – 2023. – No. 6-1. – P. 41-46.

7. Ма, Ч. Сравнение производительности алгоритмов глобального поиска с использованием набора тестовых функций / Ч. Ма // Актуальные проблемы авиации и космонавтики : Сборник материалов VII Международной научно-практической конференции, посвященной Дню космонавтики.Красноярск. 2021. – С. 79-82.

8. Ма, Ч. Стратегия рекомендаций на основе коллаборативной фильтрации для снижения ложных оценок / Ч. Ма, М. Гао // Актуальные проблемы авиации и космонавтики : Сборник материалов IX Международной научно-практической конференции, посвященной Дню космонавтики. Красноярск. 2023. – С. 74-77.

9. Ма, Z. Research and analysis of purchasing investment products based on linear programming model / Z. Ma // Актуальные проблемы авиации и космонавтики : Сборник материалов VII Международной научно-практической конференции, посвященной Дню космонавтики.Красноярск. 2021. – P. 83-86.