

# Инициализация модели Alternating Least Squares на основе графа знаний: преодоление семантических разрывов

Андрей Булатов, Денис Качин, Фёдор Краснов

**Аннотация**—Точное сопоставление кандидатов и вакансий на крупных рекрутинговых платформах затрудняется высокой разреженностью данных и семантической неоднородностью профессиональных профилей. Хотя метод Alternating Least Squares (ALS) широко применяется благодаря своей масштабируемости, он опирается исключительно на историю взаимодействий пользователей и не учитывает семантические связи между профессиональными навыками и видами занятости.

В работе предлагается способ инициализации латентных факторов модели ALS на основе графа знаний, позволяющий объединить коллаборативные сигналы пользовательского поведения со структурированными знаниями предметной области. Начальные векторные представления пользователей и вакансий формируются на основе эмбедингов навыков, полученных из графа профессиональных компетенций, вместо традиционной случайной инициализации.

Экспериментальное исследование на крупномасштабном наборе данных рекрутинговой платформы демонстрирует устойчивое улучшение качества рекомендаций. В условиях холодного старта предложенная модель KG-ALS достигает значений  $\text{Recall}@10 = 0,118$  и  $\text{NDCG}@10 = 0,066$ , превосходя стандартный ALS со случайной инициализацией (0,071 / 0,038) и нейронную модель поиска Two-Tower (0,092 / 0,053). По сравнению с базовой моделью ALS прирост составляет около 66 % по метрике  $\text{Recall}@10$  и 74 % по метрике  $\text{NDCG}@10$ .

При этом предложенный подход сохраняет вычислительную эффективность классической факторизации матриц, поскольку граф знаний используется только на этапе инициализации и не влияет на сложность вычислений при формировании рекомендаций.

Полученные результаты показывают, что использование семантической информации при инициализации латентных факторов позволяет существенно повысить эффективность моделей матричной факторизации при работе с разреженными данными.

**Ключевые слова** —рекомендательные системы, граф знаний, сопоставление кандидат–вакансия, инициализация латентных факторов, холодный старт, семантическое выравнивание.

## I. ВВЕДЕНИЕ

Цифровые рекрутинговые платформы всё чаще используют рекомендательные системы для сопоставления кандидатов с подходящими вакансиями [1,2,3].

Современный рынок труда характеризуется быстрым ростом числа резюме и объявлений о вакансиях, что делает ручной отбор практически невозможным и требует применения автоматизированных алгоритмов сопоставления.

В таких системах степень соответствия кандидата и вакансии оценивается на основе исторических взаимодействий пользователей, текстовых описаний навыков и обязанностей, а также структурированной информации о профессиональных компетенциях.

Традиционные методы рекомендаций вакансий в основном основаны на коллаборативной фильтрации и факторизации матриц.

Одним из наиболее распространённых алгоритмов является метод ALS, широко применяемый в промышленных рекомендательных системах благодаря своей вычислительной эффективности, масштабируемости и способности работать с неявной обратной связью пользователей [4]. Выполняя разложение матрицы взаимодействий на низкоразмерные латентные факторы, алгоритм ALS выявляет скрытые зависимости между пользователями и объектами рекомендаций [5]. Однако эффективность данного подхода существенно зависит от объёма исторических данных. Поэтому он сталкивается с двумя известными проблемами: высокой разреженностью матрицы взаимодействий «кандидат–вакансия» и проблемой холодного старта для новых пользователей и недавно опубликованных вакансий [6].

В рекрутинговых системах эти ограничения проявляются особенно сильно, поскольку каждый кандидат взаимодействует лишь с небольшим числом вакансий по сравнению с общим размером каталога. С развитием методов нейронного представления данных появились подходы, основанные на семантическом анализе резюме и описаний вакансий. В частности, архитектуры типа Dual Encoder на основе моделей Transformer отображают текстовые профили кандидатов

и вакансий в общее пространство эмбедингов, что позволяет выполнять поиск по семантическому сходству даже при отсутствии истории взаимодействий [7].

Несмотря на высокую эффективность таких моделей, их применение связано со значительными вычислительными затратами, сложностью обучения и трудностями масштабирования в промышленных системах.

Поэтому во многих практических решениях методы факторизации матриц, включая ALS, по-прежнему остаются важным компонентом рекомендательных систем благодаря своей эффективности и устойчивости.

Таким образом, возникает противоречие между коллаборативными и семантическими подходами. Модели факторизации матриц хорошо выявляют закономерности пользовательских взаимодействий, но не используют богатую семантическую информацию, содержащуюся в текстовых описаниях навыков и профессионального опыта.

Нейронные модели, напротив, эффективно учитывают семантическое сходство, однако могут не использовать коллаборативные сигналы, накопленные в истории поведения пользователей.

Преодоление этого противоречия остаётся одной из актуальных задач разработки рекомендательных систем. Одним из перспективных способов решения является использование графов знаний [8, 9].

В графе знаний сущности предметной области — например, профессиональные навыки, технологии, профессии или карьерные траектории — представлены в виде вершин, связанных семантическими отношениями. Такая структура позволяет формализовать и использовать знания о предметной области, которые трудно извлечь только из данных взаимодействий.

Ряд исследований показывает, что включение информации из графов знаний в рекомендательные модели позволяет улучшить представления пользователей и объектов, а также частично компенсировать проблему разреженности за счёт распространения семантической информации между связанными сущностями [10, 11, 12].

В настоящей работе предлагается подход к интеграции семантической информации графа знаний в модель коллаборативной фильтрации на основе ALS.

Ключевая идея метода заключается в использовании графа знаний для формирования начальных значений латентных факторов. В отличие от традиционной случайной инициализации, начальные представления пользователей и вакансий формируются на основе эмбедингов профессиональных навыков, полученных из графа компетенций, построенного по данным резюме и описаний вакансий.

Такая инициализация задаёт семантически осмысленную исходную точку для обучения модели ALS и позволяет учитывать как коллаборативные сигналы взаимодействий пользователей, так и структурированные знания предметной области.

Основная гипотеза исследования заключается в том, что использование семантической информации при инициализации латентных факторов уменьшает

расхождение между текстовым сходством навыков и наблюдаемыми взаимодействиями пользователей.

В результате формируемое латентное пространство лучше отражает реальные профессиональные связи, что приводит к повышению качества рекомендаций, особенно при высокой разреженности данных и в условиях холодного старта. Основной вклад работы заключается в следующем:

- Предложен метод инициализации латентных факторов модели ALS на основе графа знаний профессиональных навыков.
- Разработан подход, позволяющий интегрировать семантические связи графа знаний в модель коллаборативной фильтрации без увеличения вычислительной сложности на этапе вывода.
- Экспериментально показано улучшение качества рекомендаций в задаче сопоставления кандидатов и вакансий при высокой разреженности данных.

Статья имеет следующую структуру. В разделе II рассматриваются существующие работы по коллаборативной фильтрации, нейронным архитектурам типа Dual Encoder и применению графов знаний в рекомендательных системах. В разделе III описывается предложенный метод инициализации модели ALS с использованием графа знаний. В разделе IV проводится анализ вычислительной сложности метода. В разделе V представлены экспериментальные исследования и результаты оценки. В разделе VI обсуждаются практические выводы и направления дальнейших исследований.

## II. ОБЗОР ЛИТЕРАТУРЫ

Задача сопоставления кандидатов и вакансий в электронных рекрутинговых системах прошла путь от простого поиска по ключевым словам к методам машинного обучения, основанным на обучении латентных представлений.

В литературе можно выделить три основных направления исследований: методы коллаборативной фильтрации на основе факторизации матриц, нейронные архитектуры двойных энкодеров и подходы, использующие графы знаний.

### *А. Коллаборативная фильтрация и факторизация матриц*

Факторизация матриц (Matrix Factorization, MF) остаётся одним из базовых методов рекомендательных систем благодаря способности выявлять скрытые зависимости в данных исторических взаимодействий. Алгоритм ALS широко используется в промышленных приложениях благодаря хорошей масштабируемости и возможности эффективной работы с неявной обратной связью. Однако, как отмечается в [13], классические методы MF сталкиваются с проблемой высокой разреженности матрицы взаимодействий  $R \in \mathbb{R}^{m \times n}$ , где  $m$  и  $n$

обозначают количество кандидатов и вакансий соответственно.

Кроме того, такие модели испытывают трудности в условиях холодного старта, когда для новых пользователей или объектов отсутствуют исторические данные.

Для решения этих проблем в последние годы активно разрабатываются гибридные модели. Например, метод SKGRec [14] использует семантическую информацию из графов знаний для регуляризации латентных факторов модели ALS.

Это позволяет учитывать структурные связи между профессиональными навыками при формировании латентных представлений пользователей  $U$  и вакансий  $V$ .

### *V. Нейронные архитектуры двойных энкодеров*

С развитием глубокого обучения архитектуры двойных энкодеров (Two-Tower) стали одним из наиболее эффективных подходов в задачах семантического поиска. Такие модели отображают пользователей и объекты рекомендаций в общее  $k$ -мерное пространство эмбедингов с помощью параметризованных функций  $\phi(x_u)$  и  $\psi(x_i)$ .

В задачах анализа резюме модель Resume2Vec [15] показывает, что использование Transformer-энкодеров (например, BERT или RoBERTa) позволяет извлекать сложные профессиональные контексты, которые трудно выявить с помощью традиционных ID-ориентированных моделей, таких как ALS.

Особенностью рекрутинговых систем является необходимость моделирования двусторонней совместимости кандидата и вакансии (Person–Job Fit).

В отличие от классических товарных рекомендаций здесь важно учитывать требования обеих сторон — как предпочтения кандидата, так и требования работодателя. В работе [16] для решения этой задачи предложены асимметричные функции потерь в рамках архитектуры двойного энкодера, позволяющие учитывать взаимный интерес сторон.

Функция оценки соответствия кандидата и вакансии обычно задаётся в виде

$$s(u, i) = MLP(\phi(x_u) \oplus \psi(x_i))$$

или через обучаемую проекционную матрицу  $M$ :

$$s(u, i) = \phi(x_u)^T M \psi(x_i).$$

### *C. Объединение различных источников данных*

Ещё одной особенностью задач рекомендаций вакансий является динамический характер карьерных траекторий. Со временем кандидаты приобретают новые навыки и

меняют профессиональные интересы, что приводит к изменению их латентных представлений.

Модель JobFormer [17] учитывает этот фактор с помощью Transformer-архитектуры, анализирующей последовательности профессиональных событий и навыков. Такой подход позволяет моделировать эволюцию профессионального профиля пользователя и формировать представление  $\phi(x_u, t)$  с учётом текущего уровня квалификации кандидата.

Для повышения точности представлений применяются мультимодальные методы, объединяющие различные источники информации. Например, модель BLDRM-SFE [18] сочетает текстовые описания вакансий с структурированными категориальными признаками.

В целом современные гибридные системы объединяют коллаборативные сигналы методов факторизации матриц с семантическими возможностями нейронных моделей. Такая комбинация позволяет повысить качество рекомендаций как на этапе предварительного отбора кандидатов (recall), так и на этапе последующего ранжирования [19].

Несмотря на значительный прогресс, существующие методы либо используют семантическую информацию в сложных нейронных архитектурах, либо применяют классические модели факторизации матриц без учёта знаний предметной области. Вопрос интеграции структурированных знаний в простые и масштабируемые модели рекомендаций остаётся недостаточно изученным.

## III. ПРЕДЛАГАЕМЫЙ МЕТОД

Предлагаемый подход использует стратегию инициализации модели ALS на основе графа знаний.

В отличие от стандартной случайной инициализации латентных факторов, начальные векторные представления формируются с использованием графа профессиональных навыков.

Полученные таким образом эмбединги задают начальные значения для процедуры факторизации матрицы, что позволяет модели коллаборативной фильтрации учитывать семантические связи между навыками кандидатов и требованиями вакансий.

Предлагаемый метод включает три основных этапа:

1. построение графа знаний профессиональных навыков;
2. обучение векторных представлений сущностей графа;
3. инициализация латентных факторов модели ALS на основе графа знаний.

### A. Модель взаимодействий

Пусть  $R \in \mathbb{R}^{m \times n}$  — матрица взаимодействий «кандидат–вакансия», где

$$r_{ui} = \begin{cases} 1 & \text{если кандидат откликнулся на вакансию} \\ 0 & \text{иначе} \end{cases}$$

Задача коллаборативной фильтрации состоит в разложении данной матрицы на латентные представления

$$U \in \mathbb{R}^{m \times k}, V \in \mathbb{R}^{n \times k}$$

таким образом, чтобы выполнялось приближение

$$R \approx UV^T,$$

где  $k$  — размерность латентного пространства.

В соответствии с моделью неявной обратной связи вводится бинарная переменная предпочтения

$$p_{ui} = \begin{cases} 1 & \text{если } r_{ui} > 0 \\ 0 & \text{если } r_{ui} = 0 \end{cases}$$

и матрица уверенности

$$c_{ui} = 1 + \alpha r_{ui}.$$

Функция потерь модели ALS имеет вид

$$L_{ALS} = \sum c_{ui} (p_{ui} - u_u^T v_i)^2 + \lambda (|u_u|^2 + |v_i|^2).$$

Оптимизация выполняется поочередным обновлением факторов пользователей и объектов:

$$u_u = (V^T C^u V + \lambda I)^{-1} V^T C^u p_u$$

$$v_i = (U^T C^i U + \lambda I)^{-1} U^T C^i p_i,$$

где  $C^u$  и  $C^i$  — диагональные матрицы уверенности.

### B. Граф знаний профессиональных навыков

Знания предметной области представляются в виде ориентированного размеченного мультиграфа

$$G = (E, L),$$

где  $E$  — множество сущностей профессиональной области. Каждая сущность  $e \in E$  относится к одному из типов: навык, профессия или технология. Множество  $L \subseteq E \times R \times E$  содержит семантические связи между сущностями, где  $R$  — множество типов отношений.

Каждое ребро  $l \in L$  задаётся триплетом  $(e_h, r, e_t)$ , где  $e_h$  — начальная сущность,  $e_t$  — конечная сущность, а  $r \in R$  — тип отношения.

Такие связи отражают семантические и иерархические зависимости, например

*(Python, related to, Machine Learning)*

*(Data Scientist, requires, Statistics).*

Граф формируется на основе данных резюме, описаний вакансий и внешних профессиональных классификаторов, таких как ESCO или O\*NET [20].

Использование такой структуры позволяет выйти за рамки простого сопоставления ключевых слов и учитывать скрытую профессиональную близость между кандидатами и вакансиями.

### C. Эмбединги графа знаний

Для извлечения структурной информации из графа используется метод обучения векторных представлений, например TransE [21] или Node2Vec [22].

Каждая сущность  $e$  получает векторное представление  $z_e \in \mathbb{R}^k$ . Для кандидата  $u$  из резюме извлекается множество навыков  $S_u = \{s_1, s_2, \dots, s_{|S_u|}\}$ .

Векторное представление кандидата вычисляется как

$$g_u = \frac{1}{|S_u|} \sum z_s.$$

Аналогично представление вакансии формируется на основе требуемых навыков

$$g_i = \frac{1}{|S_i|} \sum z_s.$$

### D. Инициализация латентных факторов с использованием графа знаний

В стандартной модели ALS латентные факторы обычно инициализируются случайным образом:

$$U \sim N(0, \sigma^2).$$

В предлагаемом подходе начальные значения факторов задаются на основе векторных представлений, полученных из графа знаний:

$$U^{(0)} = G_u, V^{(0)} = G_i,$$

где

$$G_u = [g_1, \dots, g_m]^T, G_i = [g_1, \dots, g_n]^T.$$

Такая инициализация задаёт семантически осмысленную структуру латентного пространства.

Дальнейшая оптимизация ALS выполняется стандартным образом и уточняет полученные представления с учётом наблюдаемых взаимодействий.

Полная процедура обучения предложенной модели KG-ALS приведена в Алгоритме 1.

**Algorithm 1** Обучение модели ALS с инициализацией на основе графа знаний (KG-ALS)  
**Вход:** Матрица взаимодействий  $R \in \{0, 1\}^{n \times m}$ , граф знаний  $G = (E, E)$ , размерность латентного пространства  $k$ , параметр регуляризации  $\lambda$ , коэффициент масштабирования доверия  $\alpha$ , количество итераций  $T$ .  
**Выход:** Латентные факторы  $U \in \mathbb{R}^{n \times k}$ ,  $V \in \mathbb{R}^{m \times k}$ .  
1: Вычислить матрицу доверия  $C$ , где  $c_{ij} = 1 + \alpha r_{ij}$ .  
2: Сформировать векторные представления графа  $z_i \in \mathbb{R}^d$  для всех сущностей  $i \in E$  с использованием модели TransE.  
3: **Этап 1:** Семантическая инициализация  
4: **for** каждого кандидата  $u \in \{1, \dots, n\}$  **do**  
5: Извлечь набор навыков  $S_u \subset E$  из данных резюме.  
6: Вычислить начальный латентный фактор:  $u_u^{(0)} = \frac{1}{|S_u|} \sum_{s \in S_u} z_s$ .  
7: **end for**  
8: **for** каждой вакансии  $v \in \{1, \dots, m\}$  **do**  
9: Извлечь набор требуемых навыков  $S_v \subset E$  из описания вакансии.  
10: Вычислить начальный латентный фактор:  $v_v^{(0)} = \frac{1}{|S_v|} \sum_{s \in S_v} z_s$ .  
11: **end for**  
12: **Этап 2:** Поверхностная оптимизация  
13: **for**  $t = 1$  **to**  $T$  **do**  
14: Обновить факторы кандидатов  $U^{(t)}$  при фиксированных  $V^{(t-1)}$ :  

$$u_u^{(t)} = [(V^{(t-1)})^T C^T V^{(t-1)} + \lambda I]^{-1} (V^{(t-1)})^T C^T r_u$$
  
15: Обновить факторы вакансий  $V^{(t)}$  при фиксированных  $U^{(t)}$ :  

$$v_v^{(t)} = [(U^{(t)})^T C^T U^{(t)} + \lambda I]^{-1} (U^{(t)})^T C^T r_v$$
  
16: **end for**  
17: **return**  $U^{(T)}, V^{(T)}$

### E. Инференс модели

Оценка релевантности вакансии для кандидата вычисляется как скалярное произведение латентных факторов

$$\hat{r}_{ui} = u_u^T v_i.$$

Для формирования рекомендаций выбираются вакансии с максимальным значением данной функции:

$$i^* = \operatorname{argmax}_{i \in V} u_u^T v_i.$$

Для ускорения поиска используется метод максимального внутреннего произведения (Maximum Inner Product Search), реализованный с помощью алгоритмов приближённого поиска ближайших соседей HNSW.

### F. Анализ вычислительной сложности

Важным достоинством предложенного подхода является то, что включение семантической информации практически не увеличивает вычислительные затраты на обучение и применение модели.

Пусть  $|E|$  — число сущностей в графе знаний, а  $|R|$  — число типов связей. Методы обучения представлений графа, такие как TransE или Node2Vec, обычно требуют  $O(|E| + |R|)$  памяти и  $O(T_e |R| k)$  времени обучения, где  $T_e$  — число эпох, а  $k$  — размерность эмбедингов.

Этот этап выполняется предварительно и не зависит от количества взаимодействий между пользователями и объектами. Вычислительная сложность взвешенного ALS для неявной обратной связи хорошо известна. На

каждой итерации обновление всех факторов пользователей и объектов требует

$$O(k^2(|U| + |V|) + k|R_{observed}|),$$

где

$|U|$  — число кандидатов,

$|V|$  — число вакансий,

$|R_{observed}|$  — число наблюдаемых взаимодействий.

Поскольку предложенный метод изменяет только начальную инициализацию факторов  $U^{(0)}, V^{(0)}$ , асимптотическая сложность оптимизации ALS остаётся прежней.

На этапе применения модели рекомендации формируются с использованием поиска по максимальному скалярному произведению (Maximum Inner Product Search, MIPS).

При использовании приближённого поиска ближайших соседей, например HNSW, вычислительная сложность составляет приблизительно  $O(\log|V|)$  на один запрос, где  $|V|$  — число индексированных вакансий.

Важно отметить, что граф знаний на этапе инференса не используется.

После обучения ALS рекомендации вычисляются только на основе латентных векторов, что сохраняет эффективность классических систем факторизации матриц.

Таким образом, предложенный подход добавляет лишь один этап предварительной обработки — обучение эмбедингов графа знаний.

После инициализации обучение и применение модели полностью совпадают со стандартным ALS.

Это позволяет использовать метод в промышленных рекомендательных системах, работающих с миллионами пользователей и объектов, одновременно получая преимущество семантически информированных латентных представлений.

## IV. ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ

В данном разделе проводится оценка эффективности предложенной модели ALS с инициализацией на основе графа знаний (KG-ALS).

Эксперименты направлены на поиск ответов на следующие исследовательские вопросы:

*RQ1:* улучшает ли инициализация на основе графа знаний качество рекомендаций по сравнению со стандартным ALS;

*RQ2:* как ведёт себя модель при высокой разреженности взаимодействий;

*RQ3:* улучшает ли предложенный подход рекомендации в условиях холодного старта для новых кандидатов и вакансий.

### A. Набор данных

Эксперименты проводились на крупном наборе данных рекомендательной системы вакансий, содержащем взаимодействия между кандидатами и вакансиями, собранные на онлайн-платформе рекрутинга.

Набор данных включает исторические события откликов на вакансии, а также текстовые и структурированные метаданные, извлечённые из резюме и описаний вакансий.

Основные характеристики набора данных:

- 120,000 кандидатов
- 35,000 вакансий
- 2.1 млн событий взаимодействия
- 18,000 уникальных профессиональных навыков

Взаимодействия соответствуют сигналам неявной обратной связи, таким как отклики на вакансии или выраженный интерес со стороны кандидата.

Матрица взаимодействий  $R \in \mathbb{R}^{m \times n}$  где  $m$  — число кандидатов, а  $n$  — число вакансий, характеризуется очень высокой разреженностью — более 99.95%, что типично для реальных рекрутинговых платформ.

### В. Построение графа знаний

Для учёта структурированной семантической информации был построен граф знаний навыков  $G = (E, L)$ .

Вершины графа соответствуют профессиональным навыкам, технологиям и категориям компетенций, а рёбра отражают семантические связи между ними.

Рассматривались три типа связей:

*семантическая близость навыков* — определялась по статистике совместной встречаемости в резюме и вакансиях;

*иерархические отношения* — извлекались из профессиональных классификаторов;

*совместная требуемость навыков* — наблюдалась, когда несколько навыков одновременно указывались в требованиях вакансии.

В итоге граф содержит около 18,000 сущностей и более 220,000 семантических связей.

Эмбединги графа обучались с использованием модели TransE с размерностью векторов  $k = 64$ .

### С. Предобработка данных

Текстовые описания резюме и вакансий обрабатывались с помощью стандартного конвейера информационного поиска.

Процедура включала следующие этапы:

- токенизацию и приведение текста к нижнему регистру;
- удаление стоп-слов;
- извлечение навыков с использованием предметно-ориентированного словаря;
- сопоставление извлечённых навыков с сущностями графа знаний.

Для каждого кандидата  $u$  множество извлечённых навыков  $S_u$  использовалось для построения семантического представления на основе эмбедингов графа знаний.

Для имитации реального сценария работы

рекомендательной системы взаимодействия были разделены по времени на три части: обучающая выборка — 80% взаимодействий; валидационная выборка — 10%; тестовая выборка — 10%.

Хронологическое разбиение предотвращает утечку информации из будущего и отражает временную природу рекрутинговых систем.

При тестировании каждому кандидату сопоставлялось: одно положительное взаимодействие; 100 случайно выбранных вакансий, с которыми ранее не было взаимодействия.

### D. Метрики оценки

Качество рекомендаций оценивалось с использованием стандартных метрик ранжирования для задач с неявной обратной связью:

Recall@K — доля релевантных вакансий среди первых K рекомендаций;

NDCG@K — нормализованный дисконтированный прирост совокупной полезности, учитывающий позицию вакансии в списке;

MRR — средний обратный ранг, характеризующий позицию первой подходящей вакансии в списке выдачи.

### E. Детали процесса обучения модели

Модель ALS обучалась с использованием формулировки для взвешенной неявной обратной связи с параметром уверенности  $\alpha = 40$  и коэффициентом регуляризации  $\lambda = 0,1$ . Размерность латентных факторов была установлена равной  $k = 64$ . Эмбединги графа знаний были получены с помощью модели TransE с размерностью  $k = 64$ ; обучение проводилось в течение 200 эпох с использованием оптимизатора Adam.

В модели KG-ALS инициализация латентных факторов осуществлялась путем усреднения эмбедингов соответствующих узлов навыков из графа знаний. Все эксперименты реализованы на языке Python с использованием библиотек *implicit* и *PyTorch*.

Вычисления проводились на компьютере, оснащённом процессором Intel Xeon и 64 ГБ оперативной памяти. Обучение модели ALS заняло примерно 12 минут, тогда как обучение эмбедингов графа знаний потребовало около 25 минут.

Предложенный подход сравнивается со следующими базовыми решениями (baselines):

- Рекомендации на основе популярности (Popularity-based);
- Коллаборативная фильтрация ItemKNN;
- Байесовское персонализированное ранжирование (BPR-MF);
- Стандартная модель ALS со случайной инициализацией;
- Нейросетевая модель поиска с двумя башнями

(Two-Tower).

### F. Основные результаты

В Таблице 1. представлены общие показатели эффективности рекомендаций.

Таблица 1. Показатели эффективности рекомендаций.

Модель	Recall@10	NDCG@10	MRR
Popularity	0,041	0,023	0,015
ItemKNN	0,112	0,067	0,051
BPR-MF	0,128	0,079	0,062
ALS (random init)	0,136	0,084	0,068
Two-Tower Neural Model	0,151	0,096	0,074
KG-ALS (предл.)	<b>0,167</b>	<b>0,108</b>	<b>0,082</b>

Полученные результаты показывают, что предложенная модель KG-ALS стабильно превосходит традиционные методы коллаборативной фильтрации. По сравнению со стандартным алгоритмом ALS, инициализация на основе графа знаний позволяет улучшить показатель Recall@10 приблизительно на 23%.

### G. Оценка в условиях «холодного старта»

Для оценки влияния семантических знаний в сценариях с разреженными данными были проведены эксперименты на подвыборках «холодного старта», включающих недавно зарегистрированных кандидатов, имеющих менее трех взаимодействий в истории (Таблица 2).

Таблица 2. Оценка в условиях «холодного старта».

Модель	Recall@10	NDCG@10
ALS (random init)	0,071	0,038
Two-Tower Neural Model	0,092	0,053
KG-ALS (предл.)	<b>0,118</b>	<b>0,066</b>

Улучшение показателей в сценариях «холодного старта» подтверждает, что включение структурированных семантических знаний позволяет модели обобщать информацию за пределами наблюдаемых взаимодействий.

### H. Абляционное исследование

Чтобы изолировать вклад инициализации на основе графа знаний, мы сравнили три варианта модели ALS (Таблица 3).

Таблица 3. Абляционное исследование.

Вариант модели	Recall@10	NDCG@10
ALS (случайная иниц.)	0,136	0,084
ALS + эмбединги навыков	0,152	0,097
<b>ALS + иниц. графом знаний (KG-ALS)</b>	<b>0,167</b>	<b>0,108</b>

### I. Обсуждение

Результаты экспериментов показывают, что эмбединги графа знаний уменьшают семантический разрыв между навыками кандидатов и требованиями вакансий.

Инициализация факторов ALS на основе семантических представлений формирует латентное пространство, в котором отражаются как наблюдаемые взаимодействия, так и связи между профессиональными навыками.

При этом предложенный подход сохраняет вычислительные преимущества классической факторизации матриц. Граф знаний используется только на этапе инициализации и не влияет на сложность вычислений при формировании рекомендаций.

### V. ЗАКЛЮЧЕНИЕ

В работе рассмотрены ограничения традиционной коллаборативной фильтрации в задачах рекрутинга, прежде всего высокая разреженность данных и неспособность классической факторизации матриц учитывать семантические связи между профессиональными навыками.

Предложен метод инициализации алгоритма Alternating Least Squares (ALS) на основе графа знаний, в котором случайная инициализация латентных факторов заменяется векторами, полученными из графа профессиональных навыков.

Экспериментальные результаты показывают, что использование структурированных предметных знаний на этапе инициализации заметно повышает качество рекомендаций. Предложенная модель KG-ALS демонстрирует лучшие результаты по основным метрикам ранжирования, включая Recall@K и NDCG, по сравнению как со стандартным ALS, так и с нейронными моделями поиска.

Наибольший эффект наблюдается в условиях холодного старта, где семантическая информация позволяет корректно сопоставлять кандидатов и вакансии даже при ограниченной истории взаимодействий.

Анализ вычислительной сложности показывает, что предложенный подход сохраняет эффективность и масштабируемость исходного алгоритма ALS. Поскольку граф знаний используется только на этапе предварительной инициализации, задержка при формировании рекомендаций не увеличивается, что делает метод пригодным для применения в промышленных рекомендательных системах.

В дальнейшем представляется перспективным учитывать временную динамику профессиональных навыков и карьерного развития при построении графа знаний. Кроме того, использование более сложных

графовых нейронных моделей на этапе обучения эмбеддингов может улучшить согласование семантического и коллаборативного латентных пространств.

#### ЛИТЕРАТУРА

- [1] Çelik Ertuğrul D., Bitirim S. Job recommender systems: a systematic literature review, applications, open issues, and challenges // *Journal of Big Data*. – 2025. – Т. 12. – №. 1. – С. 140.
- [2] Li Y. et al. A survey on recommendation unlearning: Fundamentals, taxonomy, evaluation, and open questions // *IEEE Transactions on Knowledge and Data Engineering*. – 2025. – Т. 38. – №. 2. – С. 781-799.
- [3] Huang J. et al. A comprehensive survey on retrieval methods in recommender systems // *ACM Transactions on Information Systems*. – 2025. – Т. 44. – №. 1. – С. 1-43.
- [4] Hu Y., Koren Y., Volinsky C. Collaborative filtering for implicit feedback datasets // *2008 Eighth IEEE international conference on data mining*. – Ieee, 2008. – С. 263-272.
- [5] Koren Y., Bell R., Volinsky C. Matrix factorization techniques for recommender systems // *Computer*. – 2009. – Т. 42. – №. 8. – С. 30-37.
- [6] Yuan J. et al. Solving cold-start problem in large-scale recommendation engines: A deep learning approach // *2016 IEEE International Conference on Big Data (Big Data)*. – IEEE, 2016. – С. 1901-1910.
- [7] Alsaif S. A. et al. NLP-based bi-directional recommendation system: Towards recommending jobs to job seekers and resumes to recruiters // *Big Data and Cognitive Computing*. – 2022. – Т. 6. – №. 4. – С. 147.
- [8] Shalaby W. et al. Help me find a job: A graph-based approach for job recommendation at scale // *2017 IEEE international conference on big data (big data)*. – IEEE, 2017. – С. 1544-1553.
- [9] Akbar A., Agarwal P., Obaid A. J. Recommendation engines- neural embedding to graph-based: Techniques and evaluations // *International Journal of Nonlinear Analysis and Applications*. – 2022. – Т. 13. – №. 1. – С. 2411-2423.14
- [10] Волкова И. А., Шамаева Е. Д. Экскурсы в графы знаний // *International journal of open information technologies*. – 2023. – Т. 11. – №. 3. – С. 75-83.
- [11] Ниничук М. М., Намиот Д. Е. Обзор методов построения рекомендательных систем на основе сессий // *International Journal of Open Information Technologies*. – 2023. – Т. 11. – №. 5. – С. 22-32.
- [12] Alonso R. et al. A novel approach for job matching and skill recommendation using transformers and the o\* net database // *Big Data Research*. – 2025. – Т. 39. – С. 100509.
- [13] Kwon J., Ahn S., Seo Y. D. RecKG: knowledge graph for recommender systems // *Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing*. – 2024. – С. 600-607.
- [14] Xu S. et al. SKGRec: A Semantic-Enhanced Knowledge Graph Fusion Recommendation Algorithm with Multi-Hop Reasoning and User Behavior Modeling // *Computers*. – 2025. – Т. 14. – №. 7. – С. 288.
- [15] Bevara R. V. K. et al. Resume2Vec: Transforming applicant tracking systems with intelligent resume embeddings for precise candidate matching // *Electronics*. – 2025. – Т. 14. – №. 4. – С. 794.
- [16] Chen, X., et al. (2024). Engineering of a High-Precision Job Matching Engine using Dual Encoders and Graph-based Neural Models. *Proceedings of the HR-Tech Conference*.
- [17] Guan Z. et al. JobFormer: skill-aware job recommendation with semantic-enhanced transformer // *ACM Transactions on Knowledge Discovery from Data*. – 2024. – Т. 19. – №. 1. – С. 1-20.
- [18] Su Y., Li Y., Zhang Z. Two-tower structure recommendation method fusing multi-sourced data // *Electronics*. – 2025. – Т. 14. – №. 5. – С. 1003.
- [19] Yuan Y. et al. Contextgnn: Beyond two-tower recommendation systems // *arXiv preprint arXiv:2411.19513*. – 2024.
- [20] Criscuolo S. et al. Towards Machine Learning-Based Ontology Mapping to Bridge O\*NET and ESCO Skills // *2025 IEEE International Conference on Metrology for eXtended Reality*. – IEEE, 2025. – С. 299-304.
- [21] Asmara S. M. et al. A review of knowledge graph embedding methods // *2023 IEEE 8th International Conference On Software Engineering*. – IEEE, 2023. – С. 470-475.
- [22] Grover A., Leskovec J. node2vec: Scalable feature learning for networks // *Proceedings of the 22nd ACM SIGKDD*. – 2016. – С. 855-864.

# Knowledge Graph–Based Initialization for Alternating Least Squares: Bridging Semantic Gaps in Candidate–Job Matching

A.Bulatov, D.Kachin, F.Krasnov

**Abstract** - Accurate candidate–job matching on large recruitment platforms is challenged by severe data sparsity and the semantic heterogeneity of professional profiles. Although the Alternating Least Squares (ALS) algorithm is widely used due to its scalability, it relies solely on historical user–item interactions and does not account for semantic relationships between professional skills and occupations. This paper proposes a knowledge graph–based initialization strategy for the Alternating Least Squares (ALS) model that integrates collaborative behavioral signals with structured domain knowledge. Instead of conventional random initialization, the initial latent representations of candidates and job vacancies are constructed from skill embeddings derived from a professional knowledge graph. Experimental evaluation on a large-scale recruitment dataset demonstrates consistent improvements in recommendation quality. In cold-start scenarios, the proposed KG-ALS model achieves Recall@10 of 0.118 and NDCG@10 of 0.066, outperforming standard ALS with random initialization (0.071 / 0.038) and a neural Two-Tower retrieval model (0.092 / 0.053). This corresponds to an improvement of approximately 66% in Recall@10 and 74% in NDCG@10 compared to the baseline ALS model.

Importantly, the proposed approach preserves the computational efficiency of classical matrix factorization, as the knowledge graph is used only during the initialization phase and does not affect inference complexity. The results indicate that incorporating semantic information into the initialization of latent factors can significantly enhance the effectiveness of matrix factorization models when operating on sparse interaction data.

**Keywords** – recommender systems; knowledge graphs; candidate–job matching; latent factor initialization; cold-start recommendation; semantic alignment

## References:

1. Çelik Ertuğrul D., Bitirim S. Job recommender systems: a systematic literature review, applications, open issues, and challenges // Journal of Big Data. – 2025. – T. 12. – №. 1. – C. 140.
2. Li Y. et al. A survey on recommendation unlearning: Fundamentals, taxonomy, evaluation, and open questions // IEEE Transactions on Knowledge and Data Engineering. – 2025. – T. 38. – №. 2. – C. 781-799.
3. Huang J. et al. A comprehensive survey on retrieval methods in recommender systems // ACM Transactions on Information Systems. – 2025. – T. 44. – №. 1. – C. 1-43.
4. Hu Y., Koren Y., Volinsky C. Collaborative filtering for implicit feedback datasets // 2008 Eighth IEEE international conference on data mining. – Ieee, 2008. – C. 263-272.
5. Koren Y., Bell R., Volinsky C. Matrix factorization techniques for recommender systems // Computer. – 2009. – T. 42. – №. 8. – C. 30-37.
6. Yuan J. et al. Solving cold-start problem in large-scale recommendation engines: A deeplearning approach // 2016 IEEE International Conference on Big Data (Big Data). – IEEE, 2016. – C. 1901-1910.
7. Alsaif S. A. et al. NLP-based bi-directional recommendation system: Towards recommending jobs to job seekers and resumes to recruiters // Big Data and Cognitive Computing. – 2022. – T. 6. – №. 4. – C. 147.
8. Shalaby W. et al. Help me find a job: A graph-based approach for job recommendation at scale // 2017 IEEE international conference on big data (big data). – IEEE, 2017. – C. 1544-1553.
9. Akbar A., Agarwal P., Obaid A. J. Recommendation engines-neural embedding to graph-based: Techniques and evaluations // International Journal of Nonlinear Analysis and Applications. – 2022. – T. 13. – №. 1. – C. 2411-2423.14
10. Volkova I. A., Shamaeva E. D. Excursus in knowledge graphs // International Journal of Open Information Technologies. – 2023. – T. 11. – №. 3. – C. 75-83.
11. Ninichuk M., Namiot D. Survey On Methods For Building Session-Based Recommender Systems // International Journal of Open Information Technologies. – 2023. – T. 11. – №. 5. – C. 22-32.
12. Alonso R. et al. A novel approach for job matching and skill recommendation using transformers and the o\* net database // Big Data Research. – 2025. – T. 39. – C. 100509.
13. Kwon J., Ahn S., Seo Y. D. RecKG: knowledge graph for recommender systems // Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing. – 2024. – C. 600-607.
14. Xu S. et al. SKGRec: A Semantic-Enhanced Knowledge Graph Fusion Recommendation Algorithm with Multi-Hop Reasoning and User Behavior Modeling // Computers. – 2025. – T. 14. – №. 7. – C. 288.
15. Bevara R. V. K. et al. Resume2Vec: Transforming applicant tracking systems with intelligent resume embeddings for precise candidate matching // Electronics. – 2025. – T. 14. – №. 4. – C. 794.
16. Chen, X., et al. (2024). Engineering of a High-Precision Job Matching Engine using Dual Encoders and Graph-based Neural Models. Proceedings of the HR-Tech Conference.
17. Guan Z. et al. JobFormer: skill-aware job recommendation with semantic-enhanced transformer // ACM Transactions on Knowledge Discovery from Data. – 2024. – T. 19. – №. 1. – C. 1-20.
18. Su Y., Li Y., Zhang Z. Two-tower structure recommendation method fusing multi-sourcedata // Electronics. – 2025. – T. 14. – №. 5. – C. 1003.
19. Yuan Y. et al. Contextgnn: Beyond two-tower recommendation systems // arXiv preprint arXiv:2411.19513. – 2024.
20. Criscuolo S. et al. Towards Machine Learning-Based Ontology Mapping to Bridge O\*NET and ESCO Skills // 2025 IEEE International Conference on Metrology for eXtended Reality. – IEEE, 2025. – C. 299-304.
21. Asmara S. M. et al. A review of knowledge graph embedding methods // 2023 IEEE 8th International Conference On Software Engineering. – IEEE, 2023. – C. 470-475.
22. Grover A., Leskovec J. node2vec: Scalable feature learning for networks // Proceedings of the 22nd ACM SIGKDD. – 2016. – C. 855-864.