

*Петр ЗРЕЛОВ, Владимир КОРЕНЬКОВ,  
Николай КУТОВСКИЙ, Артем ПЕТРОСЯН,  
Борис РУМЯНЦЕВ, Роман СЕМЕНОВ, Ирина ФИЛОЗОВА*

## **АВТОМАТИЗИРОВАННАЯ СИСТЕМА МОНИТОРИНГА И АНАЛИЗА КАДРОВЫХ ПОТРЕБНОСТЕЙ ПО НОМЕНКЛАТУРЕ СПЕЦИАЛЬНОСТЕЙ ВУЗА\***

*В работе рассматривается проблема объективной оценки состояния рынка труда и подготовки выпускников, соответствующих ожиданиям работодателей, дается описание автоматизированной информационно-аналитической системы мониторинга и анализа кадровых потребностей рынка труда в выпускниках вузов (по номенклатуре специальностей высшего учебного заведения). Система предназначена для руководителей регионов, университетов, компаний, кадровых агентств. Развитие системы позволит руководителям регионов координировать открытие новых вузов или осуществлять перепрофилирование существующих в соответствии с актуальными экономическими задачами, руководителям вузов – корректировать учебные программы в соответствии с изменениями на рынке труда, компаниям – эффективно осуществлять подбор персонала и его подготовку, студентам – выбирать собственную траекторию обучения.*

**Ключевые слова:** автоматизированная информационная система, выпускник, высшее учебное заведение, иерархическая модель, машинное обучение, регион, рынок труда

Взаимодействие рынка труда и системы профессионального образования является сложным, в идеале взаимно согласованным процессом, в результате которого рынок труда покрывает свои потребности в квалифицированной рабочей силе в лице подготовленных сферой образования выпускников. Очевидно, что система профессионального образования должна отвечать потребностям рынка труда, быстро реагировать на изменения экономической ситуации. Достижение такой гибкости на практике является сложной задачей.

Кроме развития типичных форм взаимодействия учебных заведений с работодателями (участие работодателей в разработке содержания программ обучения, организация производственной практики студентов на предприятии, участие сотрудников организаций и предприятий в

---

\* Статья базируется на результатах исследований по теме «Автоматизированная система анализа, прогноза ситуации и поиска тенденций на основе больших массивов неструктурированных данных», выполнявшихся на средства ФГБОУ ВО РЭУ им. Г.В. Плеханова.

учебном процессе и трудоустройстве выпускников) не менее *важной* является *объективная оценка состояния рынка труда* в части потребности в определенных умениях и навыках, предъявляемых к потенциальной рабочей силе. Такой анализ по своей сути *находится вне сферы деятельности вузов*, а, скорее, является ответственностью структур по трудоустройству и занятости, а также местных администраций. Учебные заведения заинтересованы в таких исследованиях и должны учитывать эти данные для планирования образовательной деятельности.

Если говорить о качестве образования, с точки зрения работодателей и выпускников качество образования является низким, если уровень подготовки не соответствует требованиям производства и рынка труда. Для выпускников качество образования означает в первую очередь успешную конкуренцию. Для работодателя важна профессиональная компетентность выпускников, их способность адекватно действовать в производственной обстановке.

### ***Текущая ситуация в стандартизации высшего образования и профессиональной деятельности***

В системе государственной стандартизации программ высшего образования происходят процессы, суть которых – в отказе от жесткого нормирования содержания образования в виде набора дисциплин с фиксированной трудоемкостью и переход к рамочной регламентации структуры образовательных программ, условий их реализации и результатов освоения, представленных в форме компетенций выпускников. Вследствие этого российские организации высшего образования получают большую свободу в формировании образовательных программ, что позволяет им более адекватно реагировать на запросы рынка труда, выдерживать конкуренцию на российском и международном рынках образовательных услуг.

Федеральным законом № 122-ФЗ «О внесении изменений в Трудовой кодекс Российской Федерации и ст. 11 и 73 Федерального закона „Об образовании в Российской Федерации”», вступившим в силу 1 июля 2016 г., вводится обязанность применения работодателями профессиональных стандартов в части требований к квалификации, необходимой работнику для выполнения определенной трудовой функции, если эти требования установлены законодательством или нормативно-правовыми актами РФ. Другими словами, от профессиональных стандартов *ожидается согласование требований* к квалификациям сферы труда и сферы образования.

### ***Проблемы рынка труда и подготовки выпускников***

На данный момент для российской экономики характерно *несоответствие* количественного и качественного состава выпускников вузов потребностям рынка труда. Низкий уровень трудоустройства выпускников связан с дисбалансом спроса и предложения на рынке труда, качеством подготовки специалистов, несоответствием компетенций выпускников требованиям работодателя, а также с различными

социальными факторами. Что касается «востребованности» выпускников вузов на рынке труда, то, по данным портала «*Career.ru*», в 2014 г. в список «Топ-20» российских вузов, чьи выпускники были наиболее востребованы, попали два вуза из Санкт-Петербурга, остальные — из Москвы<sup>1</sup>. Этот факт подчеркивает остроту регионального аспекта проблемы. Анализ был проведен на основе поисковых запросов работодателей. В 2016 г. «*Career.ru*» выпустил рейтинг факультетов московских вузов по восьми профессиональным направлениям. В рейтинге использовались данные о востребованности выпускников 2015–2016 гг. факультетов/кафедр московских вузов, разместивших свои резюме на сайте «*Career.ru*» (молодежное направление «*HeadHunter*»). Оценивалась реальная востребованность выпускников посредством анализа действий соискателей/выпускников (профильность размещения) и работодателей (приглашения на собеседования, заработная плата, на которую приглашают выпускников и ее сравнение с общерыночной зарплатой) на сайте «*Career.ru*». Результаты исследований доступны по адресу<sup>2</sup>, опубликована и методология рейтинга<sup>3</sup>.

По данным исследовательской компании *MAR Consult*, изучавшей, работают ли люди по профессии, полученной в ходе обучения в вузе, по специальности не работают около половины (52%) участников исследования. Опрос проводился в Москве, Санкт-Петербурге, Екатеринбурге, Нижнем Новгороде и Самаре<sup>4</sup>.

В связи со сказанным отметим, что проблема достоверного прогноза экономического развития и подготовки соответствующих специалистов актуальна для многих стран, в т.ч. европейских, где также более востребованным становится исследование потребности в квалификациях на региональном и местном уровнях, а также на уровне отдельных предприятий. Анализ опыта деятельности по прогнозированию потребностей экономики в квалификациях в странах ЕС позволяет сделать вывод о том, что *единых системных подходов анализа рынка труда с позиций изменений требований к квалификациям рабочей силы и отражении перспективных потребностей сферы труда в содержании образовательных программ не выработано*<sup>5</sup>.

Очевидно, что эффективное прогнозирование потребностей в кадрах рынка труда возможно *только на основе объективной оценки состояния рынка*. Научно-практический интерес к данной проблеме подтверждается разработками информационно-аналитических систем, предназначенных для автоматизации сбора данных с популярных сервисов по поиску работы, и их последующему анализу с целью выявления наиболее востребованных на текущий момент специальностей и профессий<sup>6</sup>, расчета и выдачи по

<sup>1</sup> См.: URL: <https://career.ru/article/15115>.

<sup>2</sup> См.: URL: [http://mel.fm/2016/10/21/rating\\_career](http://mel.fm/2016/10/21/rating_career).

<sup>3</sup> См.: URL: [http://mel.fm/2015/10/21/metod\\_career](http://mel.fm/2015/10/21/metod_career).

<sup>4</sup> См.: *Погорелов Е.* Проблема востребованности выпускника вуза на современном рынке труда // Материалы V Международной студенческой электронной научной конференции «Студенческий научный форум». 15 февраля – 31 марта 2013 г.

<sup>5</sup> См.: *Олейникова О.Н., Муравьева А.А.* Прогнозы потребности в умениях и профессиональное образование и обучение — опыт ЕС. Центр изучения проблем профессионального образования / URL: <http://www.cvets.ru/Modules/SNA-EC.pdf>. (дата обращения 22.07.16).

<sup>6</sup> См.: *Черемисина Е.Н., Белага В.В., Самойленко Ю.И.* Информационно-образовательная среда для обучения информационным технологиям на базе Института системного анализа и управления Университета «Дубна» // Открытое образование. 2014. № 2. С. 59–65.

запросу основных индикаторов состояния рынка труда районов и региона в целом<sup>7</sup>.

Таким образом, представляется целесообразной разработка и развитие автоматизированной информационной системы (далее – АИС) мониторинга соответствия кадровых потребностей рынка и уровня подготовки выпускников.

### *Разработка АИС*

Целью реализации информационной системы мониторинга и прогноза ситуации на рынке труда и анализа кадровых потребностей является обеспечение дополнительных возможностей для выявления качественных и количественных связей между сферой образования и рынком труда.

Система разрабатывается с расчетом на широкий круг пользователей и предназначена в первую очередь для руководителей регионов, университетов, компаний, кадровых агентств. Ожидается, что реализация проекта позволит теснее связать систему образования в стране и рынок труда, предоставив возможность:

- руководителям вузов корректировать учебные программы;
- руководителям регионов открывать новые вузы или перепрофилировать существующие в соответствии с экономическими задачами субъектов Федерации;
- компаниям эффективно осуществлять подбор персонала и его подготовку.

Кроме того, предполагается, что система станет полезным инструментом для молодых специалистов, только что закончивших вуз, студентов старших курсов, начинающих искать работу по избранной специальности, а также студентов младших курсов, определяющихся со своей специализацией.

В качестве исходных данных в разработке используются ресурсы интернет-портала «*HeadHunter*»<sup>8</sup>, нормативные документы: утвержденные Федеральные государственные образовательные стандарты высшего образования (далее – ФГОС ВО) по части направлений подготовки, реализуемых в РЭУ им. Г.В. Плеханова<sup>9</sup>; реестр утвержденных профессиональных стандартов<sup>10</sup>. Насколько контент ресурса обеспечивает адекватное отражение состояния рынка труда в целом в части публикации вакансий и резюме, является предметом отдельного исследования.

### *Методическое обеспечение*

Реализация компетентного подхода к подготовке выпускников вузов регламентируется ФГОС ВО, обязательными к применению всеми имеющими государственную аккредитацию вузами, и предполагает формирование у студентов набора общекультурных, общепрофессиональных и профессиональных компетенций. Компетенция трактуется как:

<sup>7</sup> См.: Петрунина О.Е. Проектирование информационно-аналитической системы управления региональным рынком труда// Современные наукоемкие технологии. 2005. № 5. С. 75–78.

<sup>8</sup> См.: URL: <https://hh.ru/>.

<sup>9</sup> См.: URL: <http://fgosvo.ru/support/49/49/17>.

<sup>10</sup> См.: URL: <http://profstandart.rosmintrud.ru>.

- способность применять знания, умения, навыки и личностные качества для успешной деятельности в различных профессиональных ситуациях;
- интегральная норма качества образования межпредметного характера.

Профессиональные компетенции систематизированы под виды деятельности<sup>11</sup>. Для реализации ФГОС ВО по соответствующему направлению подготовки образовательное учреждение разрабатывает основную профессиональную образовательную программу (далее – ОПОП), которая включает учебный план, учебный график, рабочие программы дисциплин (модулей) и практик, фонды оценочных средств, методические материалы и другие компоненты. Планируемые результаты освоения образовательной программы (набор компетенций) указываются в общей характеристике ОПОП, в рабочие программы дисциплин (модулей) включается перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы (компетенциями).

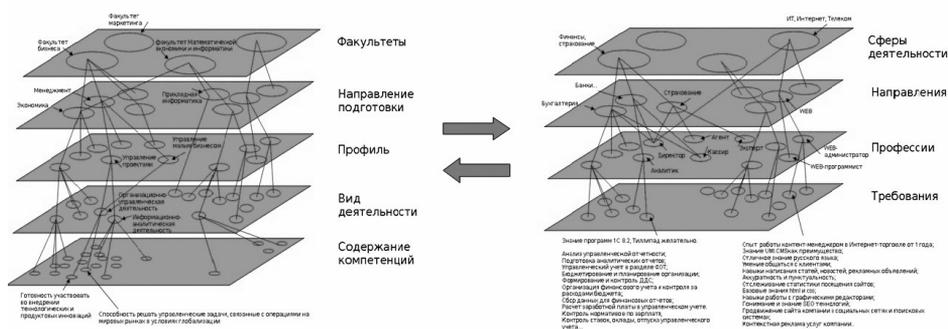
Во многих вузах ведется активная деятельность по разработке и апробации компетентностной модели выпускника (далее – КМВ) с целью комплексно описать выпускника как субъект, обладающий готовностью применения знаний, умений, навыков и личностных качеств вести продуктивную профессиональную деятельность. В КМВ обычно входят: характеристика профессиональной деятельности, требования к результатам освоения ОПОП (перечень компетенций), таблица отношений между компетенциями и учебными дисциплинами ОПОП, паспорта компетенций (совокупность требований вуза к уровню сформированности компетенции по окончании освоения ОПОП, а также развернутая характеристика требований к результатам образования в части конкретной компетенции). Это очень сложная и масштабная работа, которая еще не завершена. Таким образом, со стороны системы образования для анализа доступны формулировки содержания компетенций.

С точки зрения профессиональной деятельности можно говорить о компетентностной модели специалиста (далее – КМС) как субъекта, востребованного на рынке труда. Эту модель *описать сложнее*, т.к. работодатели не ограничены формальными рамками формулирования текстов об имеющихся вакансиях. Как уже отмечалось, ожидается, что утвержденные профессиональные стандарты могут стать связующим звеном между требованиями к квалификациям сферы труда КМС и требованиями к результатам обучения сферы образования (далее – КМВ).

Ввиду вышеизложенного на описываемом в статье этапе реализации проекта *за основу взяты упрощенные КМВ и КМС*. Выпускник описывается как субъект, обладающий набором компетенций, сформированных во время обучения по основному виду профессиональной деятельности в рамках направления подготовки по заявленному профилю. Основу КМС составляют требования, выставляемые работодателями к рабочей силе в текстах объявлений о вакансиях.

<sup>11</sup> Под компетентностью понимается уровень владения совокупностью компетенций, степень готовности к применению компетенций в профессиональной деятельности.

Идея описания предметной области в виде иерархической модели, представляющей собой ориентированный граф, вершины которого соответствуют объектам предметной области, а ребра задают отношения между ними, была заимствована из работы<sup>12</sup>. Построенные по такому принципу модели позволяют связать требования рынка и образовательные компетенции на различных уровнях. При этом образовательная модель вуза содержит 5 иерархических уровней («факультеты», «направления подготовки», «профиль», «вид деятельности», «содержание компетенций»), а модель рынка труда – 4 уровня («сферы деятельности», «направления», «профессии», «требования»). Отображение одной модели на другую происходит посредством установления связей на их нижних уровнях: для вуза это – «компетенции», для рынка труда – «требования» (см. рис. 1). Реализация связей на нижних уровнях позволяет, поднимаясь по иерархии снизу-вверх, получать связи на любом из выбранных (в зависимости от решаемой задачи) уровней. Например, отображать связи «направление подготовки» – «профессии».



**Рис. 1. Взаимное отображение между моделями системы образования и рынка труда на разных уровнях иерархий**

### Технологическое обеспечение

Сбор данных реализован в виде периодически запускаемых заданий, каждое из которых выполняет свою часть работы с данными:

- поиск новых объявлений по ключевым словам (название должности, работодателя, регион, заработная плата, список обязанностей, список требований), которые задаются в виде параметров и позволяют ограничить предметную область;
- сбор и загрузка объявлений в базу данных;
- выделение из текста объявлений значимых областей (название вакансии, региона, зарплаты, требований, обязанностей);
- подготовка текстов требований рынка труда для дальнейшего связывания.

<sup>12</sup> См.: *Гущин А.Н.* Обеспечение учебного процесса, построенного на стандартах ФГОС-3, средствами информационных технологий // *Образовательные технологии.* 2013. № 4. С. 84–89.

Система реализована с использованием свободно распространяемого программного обеспечения и может быть перенесена на любую операционную систему (*Microsoft Windows, Linux* и др.). В качестве языка разработки выбран *Python*. В качестве хранилища данных используется СУБД *MySQL* для хранения словарей, связей и файлового хранилища для сохранения исходных текстов документов. Система работает в распределенной облачной среде на основе программной платформы с открытым исходным кодом *OpenNebula*.

### **Математическое обеспечение**

Сбор и обработка данных осуществляется на основе современных методов и технологий получения информации из *web* ориентированных источников.

На следующем этапе применяются алгоритмы машинного обучения для перевода слов в векторное представление.

Далее рассчитываются вектора предложений, что позволяет выявлять смысловое сходство требований рынка труда и профессиональных компетенций высшего образования, представляющих собой не что иное, как короткие текстовые предложения. Полученные результаты используются для выявления связей на высших уровнях иерархии, описанных ранее в тексте статьи, и визуализации результатов.

### **Алгоритм связывания требований рынка и образовательных компетенций**

Моделирование семантики (смысла) слова — это *одна из ключевых проблем*, относящихся к обработке естественного языка (*Natural Language Processing*). Результаты семантического анализа используются в поисковых системах<sup>13</sup>, системах автоматического перевода<sup>14</sup> и других областях, связанных с обработкой текста на естественном языке.

На текущий момент в подходах векторного представления слов (*Word embedding*) лидирующее место занимают «предсказательные модели», основанные на использовании нейронных сетей<sup>15</sup>. Одним из главных инструментов для векторного представления слов является «*word2vec*»<sup>16</sup>.

Основной принцип работы «*word2vec*» заключается в нахождении связей между контекстами слов согласно предположению, что слова, находящиеся в похожих контекстах, имеют тенденцию обозначать похожие вещи, т.е. быть семантически близкими. Формализовать задачу, которую решает «*word2vec*», можно следующим образом: минимизировать расстояния между векторами слов, которые появляются рядом друг с другом, и максимизировать расстояния между векторами слов, которые не появляются рядом. «Рядом» в данном случае означает «в близких контекстах».

<sup>13</sup> См.: *Efrati Amir*. «Google Gives Search a Refresh». The Wall Street Journal. Retrieved July 13, 2012.

<sup>14</sup> См.: *Eva Martínez García, Cristina España-Bonet, Lluís Marquez*. Document-Level Machine Translation with Word Vector Models. Proceedings of the 18th Annual Conference of the European Association for Machine Translation (EAMT). May 2015. P. 59–66.

<sup>15</sup> См.: *Barkan Oren*. Bayesian Neural Word Embedding. arXiv:1603.06571.

<sup>16</sup> См.: *Mikolov Tomas et al*. Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781v3 [cs.CL]. 7 Sep. 2013.

Например, слова «анализ» и «исследование» часто встречаются в похожих контекстах, «word2vec» анализирует такие контексты и делает вывод, что эти слова являются близкими по смыслу. Анализ контекстов выполняется на больших корпусах текста, в нашей задаче мы использовали корпус российской Википедии и Национальный корпус русского языка. «Word2vec» содержит два типа алгоритмов — на основе модели «мешка слов» (*Continuous Bag of Words — CBOV*)<sup>17</sup> и модели «Skip-gram»<sup>18</sup>, в каждом из которых может применяться «негативное сэмплирование» (*Negative Sampling*), а также функция активации в виде нормализованной экспоненциальной функции (*Softmax*)<sup>19</sup>.

Схема работы «word2vec» такова.

Первое. Производится частотная фильтрация слов в корпусе текста.

Второе. Строится дерево Хаффмана<sup>20</sup>, применяемое для кодирования словаря — это значительно снижает вычислительную и временную сложность алгоритма.

Третье. Выбирается базовый элемент корпуса, обычно — предложение, но им может быть абзац и даже целая статья. Затем наиболее часто встречающиеся слова изымаются из анализа, что ускоряет процесс обучения алгоритма и способствует значительному увеличению качества получающейся модели.

Четвертое. Производится обход базового элемента окном, размер которого задается в качестве параметра. Под окном подразумевается максимальная дистанция между текущим и предсказываемым словом в предложении.

Пятое. Применяется нейронная сеть прямого распространения<sup>21</sup> (*Feedforward Neural Network*), которая, с одной стороны, использует модель «мешка слов», предсказывая слово по контексту, — на рисунке 2 (слева) входной слой «input» представляет собой набор слов из окна  $\{w(t-2) \dots w(t+2)\}$ , т.е. контекст, а выходной слой «output»  $w(t)$  — слово, которое нужно предсказать, вычисляя сумму входного вектора. Результат сравнивается с выходным вектором, и процедура повторяется до тех пор, пока не будет достигнуто достаточное соответствие. С другой стороны, модель «Skip-gram» предсказывает контекст по слову, являясь «зеркальным» отражением алгоритма «мешка слов» — на рисунке 2 (справа). Результатом является отображение слов в  $n$ -мерное метрическое пространство ( $n$  — параметр модели).

Существуют попытки создать предсказательную модель для перевода документа в векторное пространство<sup>22</sup>. Однако задача сравнения коротких предложений на смысловую схожесть обладает определенной спецификой, и использование существующих моделей по переводу слов или документов в векторное пространство без модификаций *дает неудовлетворительный результат*.

<sup>17</sup> См.: Harris Zellig. Distributional Structure. Word 10 (2–3): August 1954. P. 146–162.

<sup>18</sup> См.: Broder Andrei Z., Glassman Steven C., Manasse Mark S., Zweig Geoffrey. Syntactic clustering of the web. Computer Networks and ISDN Systems 29 (8): 1157–1166. doi:10.1016/s0169-7552(97)00031-7.

<sup>19</sup> См.: Christopher M. Bishop. Pattern Recognition and Machine Learning. Springer. 2006.

<sup>20</sup> См.: Huffman D. A Method for the Construction of Minimum-Redundancy Codes (PDF). Proceedings of the IRE. 40 (9): 1098–1101. doi:10.1109/JRPROC.1952.273898.

<sup>21</sup> См.: Розенблатт Ф. Принципы нейродинамики: Перцептроны и теория механизмов мозга. М.: Мир, 1965.

<sup>22</sup> См.: Le Quoc et al. Distributed Representations of Sentences and Documents. arXiv:1405.4053.

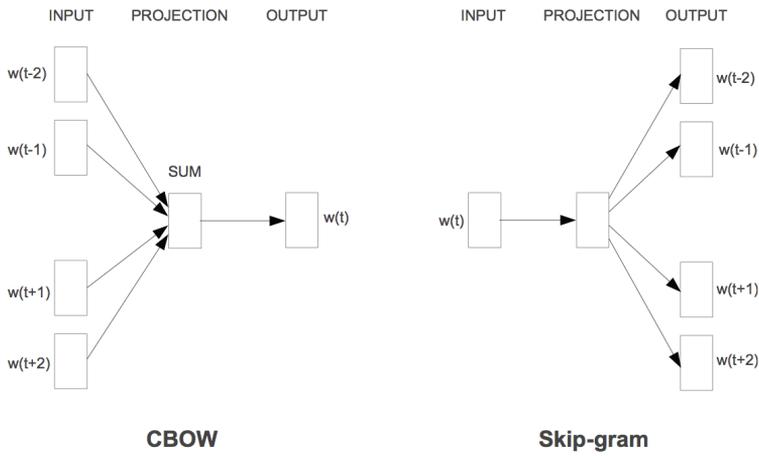


Рис. 2. Схема работы алгоритмов «мешка слов» (CBOW) и «n-грамм» (Skip-gram)

Поскольку тексты формулировок образовательных компетенций, так же как и формулировки требований в объявлениях о вакансиях, содержат в среднем около 10 слов, то в основе аналитической части системы лежит задача вычисления семантической близости двух коротких предложений. Авторами был разработан алгоритм перевода предложений в векторное пространство, основанный на «word2vec».

Таким образом, каждому слову в соответствие ставится вектор размерностью  $n$ , которая характеризует слово и влияет на точность модели. Метрическое пространство отображений слов принято называть семантическим. В качестве примера на рисунке 3 представлены проекции векторов на плоскость. Близкие по смыслу слова находятся рядом и образуют некоторые смысловые кластеры.

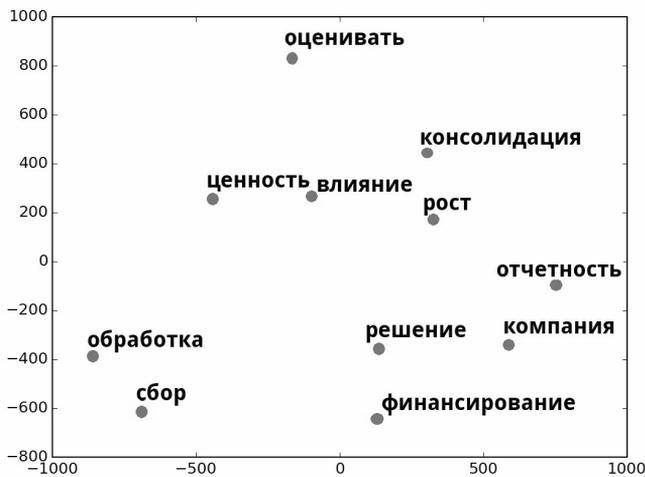


Рис. 3. Распределение векторов слов в проекции на плоскость

Векторное представление позволяет вычислять «похожесть» слов на основе расчета косинусного расстояния.

Так, для двух слов  $w_1$  и  $w_2$ , представленных в виде векторов  $\vec{V}(w_1)$  и  $\vec{V}(w_2)$ , семантическая близость рассчитывается по формуле:

$$\cos(\vec{V}(w_1), \vec{V}(w_2)) = \frac{\vec{V}(w_1) \times \vec{V}(w_2)}{|\vec{V}(w_1)| \cdot |\vec{V}(w_2)|} \quad (1)$$

По аналогии с вычислением похожести слов рассчитывается семантическая близость компетенций и требований, которые представляют собой короткие предложения, имеющие в своем составе в среднем 10 слов. Расчет вектора описанных предложений  $\vec{V}(s)$ , где  $s = \{w_1, w_2, \dots, w_k\}$ , определяется как среднее взвешенное от векторов слов, из которых оно состоит:

$$\vec{v}(s) = \frac{\sum_{i=1}^k p_i \cdot \vec{v}(w_i)}{\sum_{i=1}^k p_i}, \quad (2)$$

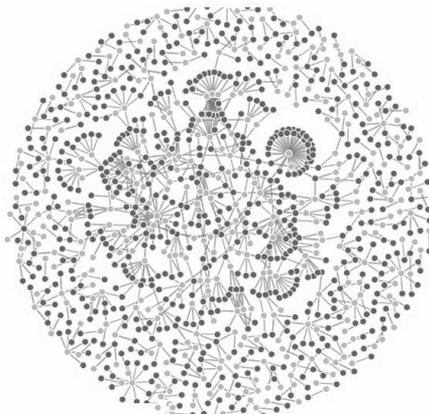
где

$p_i$  – вес слова, который рассчитывается как отношение частоты употребления слова к размерности лексикона выбранного уровня иерархии на стороне системы образования или рынка труда,

$k$  – количество слов в предложении.

После чего рассчитывается семантическая близость предложений по формуле, приведенной выше. Стоит отметить, что слова, не имеющие смысловой нагрузки (союзы, частицы, предлоги, местоимения и т.д.), не участвуют в формировании вектора предложения.

Одним из методов визуализации результатов сравнения является построение взвешенного графа (см. рис. 4), отражающего связи между отдельными компетенциями и требованиями рынка труда. Вершины соответствуют образовательным компетенциям (обозначены серым цветом) и требованиям рынка труда (обозначены черным), ребра графа отражают наличие и силу связи между ними.



**Рис. 4.** Взвешенный граф, отражающий связи между образовательными компетенциями и требованиями рынка труда, накопленными в базе данных системы

Вычислительная сложность алгоритма достаточно велика, и время построения матрицы смежности после применения процедуры распараллеливания на сервере с 48 ядрами (96 потоков) *Intel Xeon E7-8850 v2* частотой 2,30 ГГц и 256 ГБ ОЗУ может быть оценено примерно в 4 часа (для данных, собранных на настоящий момент в базе системы).

### ***Перспективы развития алгоритма***

В силу того, что сравниваемые предложения *имеют узкую направленность*, а российская Википедия и Национальный корпус русского языка охватывают огромное число сфер и видов деятельности, модель получается довольно размытой относительно задачи. Главным образом это проявляется в отсутствии векторов для некоторых слов либо их вариаций. Для частичной ликвидации подобного эффекта принято решение сделать модель двухуровневой: второй уровень представляет собой тот же алгоритм сравнения предложений, что описан выше, однако он работает не со словами, а с основами слов, т.е. с их неизменными частями. Авторы предполагают, что накопление базы вакансий позволит сформировать уникальный корпус, учитывающий специальную терминологию рынка труда, который затем будет использоваться в обучении моделей.

Также стоит отметить, что подтверждение адекватности результатов сравнения теоретически возможно с использованием вмешательства экспертов, однако объемы полученных результатов свидетельствуют о фактической невозможности полноценной проверки в разумные сроки. Поэтому авторами разрабатываются методы, позволяющие верифицировать работу данной модели.

### ***Описание АИС***

Реализованный прототип автоматизированной информационной системы представляет собой *web* ориентированное приложение с интуитивно-понятным пользовательским интерфейсом, обеспечивает надежное хранение данных.

Система построена по модульному принципу и включает, во-первых, модуль сбора текстовых данных (функционирующий в автоматическом режиме с использованием открытых источников, в качестве которых выступает интернет-портал, аккумулирующий информацию кадровых агентств).

Во-вторых, модуль загрузки и хранения данных, состоящий из базы данных и распределенного хранилища данных (обеспечивающего репликацию и архивирование).

В-третьих, модуль автоматической обработки, выполняющий подготовку информации для анализа, автоматическое связывание требований и компетенций, машинное обучение.

В-четвертых, модуль генерации и отображения отчетов.

В-пятых, пользовательский интерфейс.

На *рисунке 5* представлен общий вид главного окна системы.



Рис. 5. Главное окно web-интерфейса прототипа системы

Генератор отчетов позволяет представлять результаты как в текстовом, так и в графическом виде. Также можно проводить количественный анализ, например, выявлять наиболее популярные требования к кандидату для той или иной профессии, или выявлять профили направления подготовки вуза, соответствующие наибольшему количеству вакансий на рынке и т.д. В качестве примера работы системы на *рисунке 6* представлен результат анализа соответствия профилей подготовки, а на *рисунке 7* — направлений подготовки требованиям рынка труда в РЭУ им. Г.В. Плеханова. Другим примером является анализ востребованности рынком труда выпускников РЭУ в городах России. Поскольку карта связанности содержит информацию о связи между компетенциями и требованиями, можно установить связь «компетенция — вакансия». Исходя из полученного распределения (см. *рис. 8*), можно сделать заключение, в каких городах выпускникам РЭУ будет проще найти работу.



Рис. 6. Круговая диаграмма, отражающая востребованность профилей подготовки в РЭУ им. Г.В. Плеханова на рынке труда (на январь 2016 г.)



Рис. 7. Круговая диаграмма, отражающая востребованность направлений подготовки в РЭУ им. Г.В. Плеханова на рынке труда (на январь 2016 г.)

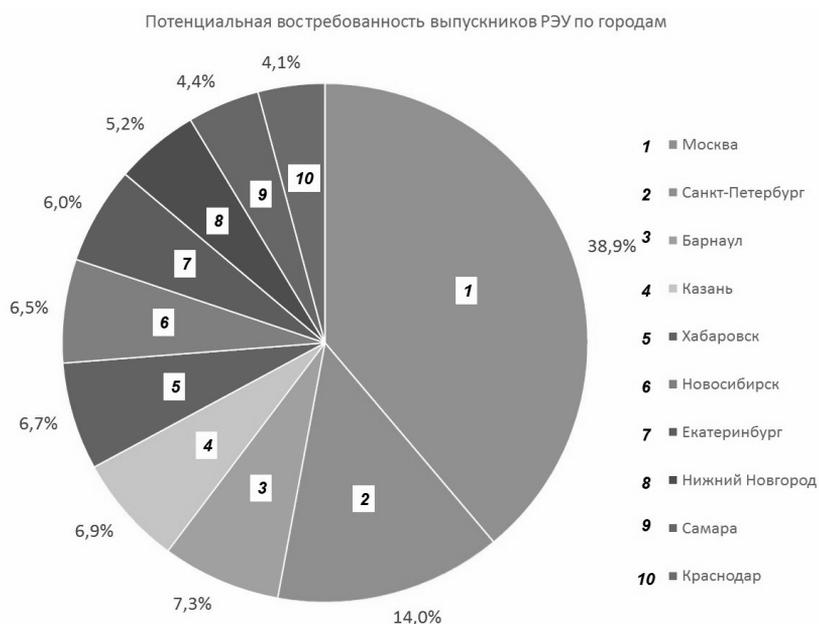


Рис. 8. Круговая диаграмма, отражающая потенциальную востребованность выпускников РЭУ в различных городах РФ (на октябрь 2016 г.)

\* \* \*

В рамках реализации проекта создан прототип автоматизированной информационной системы мониторинга и анализа кадровых потребностей регионов по номенклатуре специальностей вуза. Прототип разработан для решения задач РЭУ им. Г.В. Плеханова, в т.ч.

для включения в состав программных и технологических решений Ситуационного центра социально-экономического развития России и субъектов Федерации.

С помощью этой системы, в результате анализа постоянно обновляющихся больших массивов данных, можно устанавливать, насколько программы обучения высшего образования соответствуют текущим ожиданиям рынка, прогнозировать изменение этих ожиданий и автоматически выдавать рекомендации для корректировки учебных программ с целью наиболее точного соответствия этим ожиданиям. Развитие и адаптация системы могут производиться в соответствии с требованиями заказчика в зависимости от специфики задачи — особенностей региона, вуза и пр. Полагаем, что созданная система, а также алгоритмы и принципы, на которых она построена, в дальнейшем могут быть использованы и для решения более широкого класса актуальных проблем. Для этого система может быть перестроена в зависимости от особенностей в постановке задачи и типа входных данных.