

## ПАТЕНТНЫЙ ПОИСК НА ОСНОВЕ СЕМАНТИЧЕСКОГО АНАЛИЗА

**Яворская В.А.**

ФБГОУ ВПО «Ижевский государственный технический университет имени  
М.Т.Калашникова»,  
Россия, 426069, Удмуртская республика, город Ижевск, Студенческая, 7  
+7 (963) 543-38-63 [valeriya.yavorskaya@gmail.com](mailto:valeriya.yavorskaya@gmail.com)

В настоящее время остро стоит проблема быстрого принятия решения при патентной экспертизе. Одним из аспектов экспертизы является поиск по уровню техники: поиск патентов, способных опровергнуть заявку. Анализ нового патента по уровню техники проводится, в основном, с использованием программных средств, в основе функционирования которых лежит поиск разной степени сложности по ключевым словам. Такой тип поиска приводит к большим временным затратам как со стороны работы программного средства, так и со стороны работы оператора, которому зачастую приходится самостоятельно обработать большой объём информации о патентах, сходных с запрашиваемым. Используемые при этом программные средства являются по существу поисковыми и не имеют соответствующего функционала принятия решений.

Одним из вариантов решения данной проблемы является разработка программного средства, функционирующего на основании семантического поиска, которое бы не только подбирало похожие патенты по уровню техники, но и принимало решение о принятии или отклонении патентной заявки.

Задача семантического поиска давно рассматривается учёными, предлагаются различные методы её решения, в частности, использование модели вероятностного латентно-семантического анализа (pLSA), а также латентного размещения Дирихле (LDA). Вторая модель получила большее распространение, поскольку позволяет вносить новые тематики без полной перестройки существующей сети.

В данной работе рассматриваются расширения модели pLSA, которые могут быть использованы в качестве замены LDA для повышения скорости обработки больших объёмов данных (здесь — баз патентов) без потери корректности внесения новых тематик. Это позволит быстрее и качественнее производить поиск по уровню техники для патентных заявок.

### **Литература.**

1. *Blei David M., Ng Andrew Y., Jordan Michael I.* Latent Dirichlet Allocation // Journal Of Machine Learning Research 3 2003 p.993-1022
2. *Hoffman Thomas* Unsupervised Learning By Probablistic Latent Semantic Analysis // Machine Learning 2001 42 p.177-196
3. *Воронцов К.В., Потапенко А.А.* Регуляризация, робастность и разреженность вероятностных тематических моделей // Компьютерные исследования и моделирование Т.4 2012 №4 с.693-706