

УДК 347.77

ВЛИЯНИЕ ИСКАЖЕНИЙ И ШУМОВ ПРИ ПОСТРОЕНИИ ПАТЕНТНОГО ЛАНДШАФТА НА КАЧЕСТВО УПРАВЛЕНИЯ ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТЬЮ

Д.А. Журавлёв, С.В. Мурашова, А.И. Литвинов

Федеральное государственное автономное образовательное учреждение высшего образования "Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики"

Аннотация: В статье оценена величина ошибки от искажений и шумов патентного поиска, проявляющихся на этапе поиска и отбора требуемых документов по заданному техническому результату. Новизной представленного материала является рассмотрение процесса поиска на нескольких этапах с оценкой не только ошибки, но и затрачиваемого временного ресурса.

Ключевые слова: патентный ландшафт, искажения и шумы патентный поиск, патентная информация, единичный запрос, множественный запрос, поисковая система.

EFFECTS OF DISTORTION AND NOISE WHEN BUILDING A PATENT LANDSCAPE ON THE INTELLECTUAL PROPERTY MANAGEMENT QUALITY

D. Zhuravlev, S. Murashova, A. Litvinov

ITMO University

Abstract: The article assesses the magnitude of the error from the distortion and noise of patent search, manifested at the stage of search and selection of the required documents for a given technical result. The novelty of the presented material is the consideration of the search process at several stages with an assessment of not only the error, but also the time spent.

Key words: patent landscape, distortions and noises patent search, patent information, single request, multiple request, search system.

Введение. Патентный ландшафт является одним из базовых элементов управления интеллектуальной собственностью при определении и уточнении научно-технологических приоритетов на основании анализа патентной информации [1, 6]. Рассмотренные в существующей литературе проблемы построения патентных ландшафтов связаны в основном с необходимостью обработки большого объема патентной и дополнительной информации, доступом к базам данных и уровнем подготовки специалиста, подготавливающего патентный ландшафт [2, 6]. В тоже время одной из основных проблем построения патентных ландшафтов, проявляющейся на этапе сбора данных, являются искажения и шумы патентного поиска ($\epsilon_{\text{ип}}$). Эта проблема вызвана сложностью

работы с большими массивами неструктурированных данных, как при обработке патентной информации, где сосредоточен основной объем сведений (70...80%), так и при работе с дополнительными источниками, например, журналами на которые приходится всего лишь 20...30% [3, 4].

Ошибка от искажений и шумов патентного поиска.

Ошибка от искажений и шумов патентного поиска $\epsilon_{\text{ип}}$ может быть при единичном запросе без изменения ключевых слов и при множественном запросе, изменении ключевых слов на слова синонимы.

Величина ошибка от искажений и шумов патентного поиска при единичном запросе вычисляется

$$\varepsilon_{\text{ип}} = \left(\frac{N_i - N_j}{N_i} \right) 100 \% \quad (1)$$

где N_i – общее количество документов при поиске; N_j – количество документов «отсеянных» при обработке, не относящихся к тематике поиска или техническому результату.

Так как требуемые документы по тематике поиска на каждом его этапе входят в

общее количество найденных документов, то в формуле (1) на каждом этапе поиска необходимо вычитать этот процент из величины ошибки. Тогда величина ошибка от искажений и шумов патентного поиска при единичном запросе будет вычисляться как

$$\varepsilon_{\text{ип}} = \left(\frac{N_i - N_j}{N_i} 100 \% \right) - \left(100 \% - \left(\frac{N_i - N_g}{N_i} 100 \% \right) \right), \quad (2)$$

где N_g – количество документов относящихся к тематике поиска или техническому результату.

Величина ошибка от искажений и шумов патентного поиска при множественном запросе вычисляется

$$\varepsilon_{\text{ип}} = \left(\frac{N_y - N_z}{N_y} 100 \% \right) - \left(100 \% - \left(\frac{N_y - N_m}{N_y} 100 \% \right) \right), \quad (3)$$

где $N_y = \sum_{i=1}^k N_i$ – количество несовпадающих документов при каждой смене ключевых слов из общего количества найденных документов – N_a , $N_y \in N_a$;

N_z – количество несовпадающих документов при каждой смене ключевых слов «отсеянных» при обработке, не относящихся к тематике поиска или техническому результату, $N_z \in N_y$; N_m – общее количество несовпадающих документов, при каждой смене ключевых слов, относящихся к тематике поиска или техническому результату, $N_m \in N_y$ [5].

Пример проявления искажений и шумов патентного поиска.

На этапе сбора данных задается тема исследования – способ изготовления одномодового оптического волокна. Технический результат – изготовление одномодового оптического волокна. Цель исследования – оценить величину ошибки от искажений и шумов патентного поиска $\varepsilon_{\text{ип}}$ при единичном запросе без изменения ключевых

слов и при множественном запросе, при изменении ключевых слов на слова синонимы. Ограничения и допущения: использовалась одна база данных – ФИПС; уровень подготовки специалиста (патентоведа) – средний.

Оценка ошибки патентного поиска от искажений и шумов патентного поиска при единичных запросах без изменения ключевых слов при первоначальном (1 этап), уточняющем (2 этап) и основательном поиске и анализе (3 этап).

Результаты работы поисковой системы ФИПС по поиску патентных документов на тему «способ изготовления одномодового оптического волокна».

Первоначальный поиск (ПП). Выявлено 162 патентных документов. Уточняющий поиск и анализ (УПиА) направленный на отбор документов, связанных только с оптическими волокнами, позволил сократить их количество до 75. Анализ был осуществлен за 45 минут. Основательный поиск и анализ (ОПиА) направленный на отбор документов, связанных только с способом изготовления оптических волокон, позволил сократить их количество до 19 из них 9 документов посвящено способам

изготовления одномодовых оптических волокон и 10 патентных документов посвящено способам изготовления многомодовых оптических волокон, причем в данном случае эти 10 документов относятся так же к шумам патентного поиска. Анализ был осуществлен за 2 часа. Величина ошибка от искажений и шумов патентного поиска при единичном запросе вычисляется по формуле (2):

при первоначальном поиске $\varepsilon_{\text{ш}} \rightarrow 94,45$ % за $t = 0,02495 \approx 0,025$ мин;

при уточняющем поиске и анализе $\varepsilon_{\text{ш}} \rightarrow 40,74$ % за $t = 45$ мин;

при основательном поиске и анализе $\varepsilon_{\text{ш}} \rightarrow 0$ % за $t = 120$ мин.

Таким образом, при единичных запросах без изменения ключевых слов при основательном поиске и анализе можно обеспечить искажения и шумы патентного поиска 0 % за время 165 мин. Оценка ошибки патентного поиска от искажений и шумов патентного поиска при изменении ключевых слов на слова синонимы, гипонимы и гиперонимы к слову оптическое волокно при аналогичном, ранее рассмотренном, делении на этапы: первоначальный, уточняющий и основательный поиск и анализ.

Результаты работы поисковой системы ФИПС по поиску патентных документов по тематике: № 1 – «изготовление одномодового оптического волокна» (226 документов); № 2 –

«изготовление одномодовой оптической заготовки» (86 документов); № 3 – «изготовление одномодового оптического световода» (141 документов). Расчет искажений и шумов патентного поиска при изменении ключевых слов. Значения для каждой области запроса при расчетах представлено соответствующим номером.

При ПП выявлено № 1 – 226; № 2 – 86; № 3 – 141 патентных документов. Во-время УПиА количество документов сокращено до № 1 – 79; № 2 – 28; № 3 – 43 штук. Анализ был осуществлен за 1 час 55 минут; 40 мин; 1 час 10 мин, соответственно. ОПИА позволил сократить их количество до № 1 – 21; № 2 – 16; № 3 – 19 из них № 1 – 11; № 2 – 6; № 3 – 9 документов посвящено способам изготовления одномодовых оптических волокон и № 1 – 10; № 2 – 10; № 3 – 10 патентных документов посвящено способам изготовления многомодовых оптических волокон, причем патентные документы посвященные способам изготовления многомодовых оптических волокон относятся так же к шумам патентного поиска. Анализ был осуществлен за 1 час 5 минут; 30 мин; 44 мин, соответственно. Величина ошибка от искажений и шумов патентного поиска при изменении ключевых слов для каждого вышеуказанного запроса (единичный запрос) вычисляется по формуле (2) и представлена в табл. 1.

Таблица 1

Основная область поиска	Искажения и шумы патентного поиска		
	При первоначальном поиске, %	При уточняющем поиске и анализе, %	При основательном поиске и анализе, %
Изготовление одномодового оптического волокна	95,13	28,73	0
Изготовление одномодовой оптической заготовки	93,02	25,57	0
Изготовление одномодового оптического световода	93,62	24,12	0

Время, затраченное на поиск аналогично расчетам при оценке ошибки патентного поиска от искажений и шумов патентного поиска при единичных запросах без изменения ключевых слов и ввиду малой величины, по сравнению с УПиА и ОПиА, приравнено к нулю. Общее время поиска 364 мин. Величина ошибка от искажений и шумов патентного поиска при множественном запросе вычисляется по формуле (3).

Общее количество найденных документов $N_a = 615$, количество несовпадающих документов при каждой смене ключевых слов $N_y = 263$, количество несовпадающих документов при каждой смене ключевых слов «отсеянных» при обработке, не относящихся к тематике поиска или техническому результату N_z при: ПП выбрано 263 патентных документов. Анализ был осуществлен за 4 часа 36 минут; УПиА выбрано 87 документов. Анализ был осуществлен за 1 час 45 минут; ОПиА выбрано 26 документов из них 12 связанных с заданной тематикой - способ изготовления одномодовых оптических волокон. Анализ был осуществлен за 1 час 36 минут. Из 12 найденных документов 9 документов было найдено в единичном поиске по запросу «способ изготовления одномодового оптического волокна», а 3 документа было найдено новых при множественном запросе, которые не попали в единичный запрос. Общее время поиска, анализа и обработки составило 477 мин.

Оценка величины ошибки и времени при патентном анализе с увеличением слов при запросе с 4-х ранее рассмотренных (1 слово – при единичном и 3 слова синонима при множественном) до 10 и 16 слов (увеличение на 6 слова) при дополнительном 2-ом и 3-ем множественном поиске, соответственно.

Результаты работы поисковой системы ФИПС по поиску патентных документов по тематике на втором множественном поиске: № 1 – «изготовление одномодового волновода» (112 документов); № 2 – «способ изготовления одномодового волновода» (75 документов); № 3 – «изготовление одномодовой оптической нити» (44 документа); № 4 – «способ изготовления одномодовой оптической нити» (27 документа); № 5 – «изготовление

одномодового оптического стекловолокна» (28 документа) и № 6 – «способ изготовления одномодового оптического стекловолокна» (22 документа). Значения для каждой области запроса при расчетах представлено соответствующим номером.

ПП – выявлено № 1 – 112; № 2 – 75; № 3 – 44; № 4 – 27; № 5 – 28; № 6 – 22 патентных документов. УПиА – позволил сократить их количество до № 1 – 41; № 2 – 19; № 3 – 23; № 4 – 12; № 5 – 17; № 6 – 9 патентных документов. Анализ был осуществлен за 1 час 44 минуты; 55 мин; 36 мин; 18 мин; 20 мин; 17 мин, соответственно. ОПиА позволил сократить их количество до № 1 – 13; № 2 – 13; № 3 – 7; № 4 – 7; № 5 – 4; № 6 – 4 из них № 1 – 6; № 2 – 6; № 3 – 2; № 4 – 2; № 5 – 2; № 6 – 2 документов посвящено способам изготовления одномодовых оптических волокон и № 1 – 7; № 2 – 7; № 3 – 5; № 4 – 5; № 5 – 2; № 6 – 2 патентных документов посвящено способам изготовления многомодовых оптических волокон. Анализ был осуществлен за 1 час 10 мин; 44 мин; 23 мин; 20 мин; 24 мин; 15 мин, соответственно. На этапах УПиА и ОПиА поиск и анализ для каждого поискового запроса осуществлялся последовательным способом. Общее время поиска, анализа и обработки составило 446 мин. Величина ошибка от искажений и шумов патентного поиска при изменении ключевых слов для каждого вышеуказанного запроса (единичный запрос) вычисляется по формуле (2) и представлена в табл. 2. Общее количество найденных документов $N_a = 308$,

количество несовпадающих документов при каждой смене ключевых слов $N_y = 151$, количество несовпадающих документов при каждой смене ключевых слов «отсеянных» при обработке, не относящихся к тематике поиска или техническому результату N_z при: ПП выбрано 151 патентных документов.

Анализ был осуществлен за 3 часа 15 минут; УПиА выбрано 42 документа. Анализ был осуществлен за 1 час 10 минут; ОПиА выбрано 19 документов из них 8 связанных с заданной тематикой - способ изготовления одномодовых оптических волокон. Анализ был осуществлен за 55 минут. Общее время поиска, анализа и обработки составило 320 мин.

Анализ результатов по оценке ошибки патентного поиска от искажений и шумов патентного поиска показывает, что при единичных запросах без изменения ключевых слов величина ошибки составила 0 %, при увеличении области запроса до 3 синонимов величина ошибки относительно единичного поиска увеличилась на 25 % (найдено 3 новых патентных документов относительно

единичного запроса - 9 документов). При втором множественном поиске было найдено еще 4 документа не попавших ни в единичный запрос, ни в первый множественный. Таким образом ошибка патентного поиска составляет: относительно первого множественного – увеличилась на 25 %, а относительно единичного – увеличилась на 43,75 %.

Таблица 2

Основная область поиска	Искажения и шумы патентного поиска		
	При первоначальном поиске, %	При уточняющем поиске и анализе, %	При основательном поиске и анализе, %
Изготовление одномодового волновода	95,22	31,3	0
Способ изготовления одномодового волновода	92	17,3	0
Изготовление одномодовой оптической нити	95,46	47,73	0
Способ изготовления одномодовой оптической нити	92,6	37	0
Изготовление одномодового оптического стекловолокна	92,85	53,6	0
Способ изготовления одномодового оптического стекловолокна	90,9	31,8	0

Результаты работы поисковой системы ФИПС по поиску патентных документов по тематике на третьем множественном поиске: № 1 – «изготовление одномодового оптоволокна» (34 документа); № 2 – «способ изготовления одномодового оптоволокна» (25 документов); № 3 – «изготовление оптического одномодового световолокна» (6 документа); № 4 – «способ изготовления оптического

одномодового световолокна» (3 документа); № 5 – «изготовление оптического одномодового оптиковолокна» (0 документов) и № 6 – «способ изготовления оптического одномодового оптиковолокна» (0 документов).ПП –выявлено № 1 – 34; № 2 – 25; № 3 – 6; № 4 – 3; № 5 и № 6 – 0 патентных документов. УПиА позволил сократить их количество до № 1 – 12; № 2 – 12; № 3 – 1; № 4

– 1; № 5 и № 6 – 0 патентных документов, поэтому в дальнейшем в УПиА и ОПиА не учитываются. Анализ был осуществлен за 30 мин; 25 мин; 9 мин; 3 мин; 0 мин; 0 мин, соответственно. Основательный поиск и анализ (ОПиА) направленный на отбор документов, связанных только с способом изготовления оптических волокон, позволил сократить их количество до № 1 – 5; № 2 – 5; № 3 и № 4 – 0 из них № 1 – 2; № 2 – 2 документов посвящено способам изготовления одномодовых оптических волокон и № 1 – 3; № 2 – 3 патентных документов посвящено способам

изготовления многомодовых оптических волокон, причем патентные документы посвященные способам изготовления многомодовых оптических волокон относятся так же к шумам патентного поиска. Анализ был осуществлен за 20 мин; 12 мин; 3 мин; 2 мин; 0 мин; 0 мин, соответственно. Величина ошибка от искажений и шумов патентного поиска при изменении ключевых слов для каждого вышеуказанного запроса (единичный запрос) вычисляется по формуле (2) и представлена в табл. 3.

Таблица 3

Основная область поиска	Искажения и шумы патентного поиска		
	При первоначальном поиске, %	При уточняющем поиске и анализе, %	При основательном поиске и анализе, %
Изготовление одномодового оптоволокна	94,12	29,4	0
Способ изготовления одномодового оптоволокна	92	42,12	0
Изготовление оптического одномодового световолокна	100	100	100
Способ изготовления оптического одномодового световолокна	100	100	100
Изготовление оптического одномодового оптиковолокна	–	–	–
Способ изготовления оптического одномодового оптиковолокна	–	–	–

Общее количество найденных документов $N_a = 68$, количество несовпадающих документов при каждой смене ключевых слов $N_y = 39$, количество несовпадающих документов при каждой смене ключевых слов «отсеянных» при обработке, не относящихся к тематике поиска или техническому результату N_z при: ПП выбрано 39 патентных документов. Анализ был осуществлен за 1 час 38 минут;

УПиА выбрано 13 документов. Анализ был осуществлен за 48 минут; ОПиА выбрано 5 документов из них 2 связанных с заданной тематикой - способ изготовления одномодовых оптических волокон. Анализ был осуществлен за 15 минут. Общее время поиска, анализа и обработки составило 161 мин.

Анализ результатов по оценке ошибки патентного поиска от искажений и шумов патентного поиска показывает, что при третьем

множественном поиске был обнаружен только один новый патентный документ, который не попал ни в первый единичный запрос, ни в последующие первый и второй множественные

запросы.

Зависимость величины ошибки патентного поиска от временных затрат при патентном поиске показана на рисунок 1.

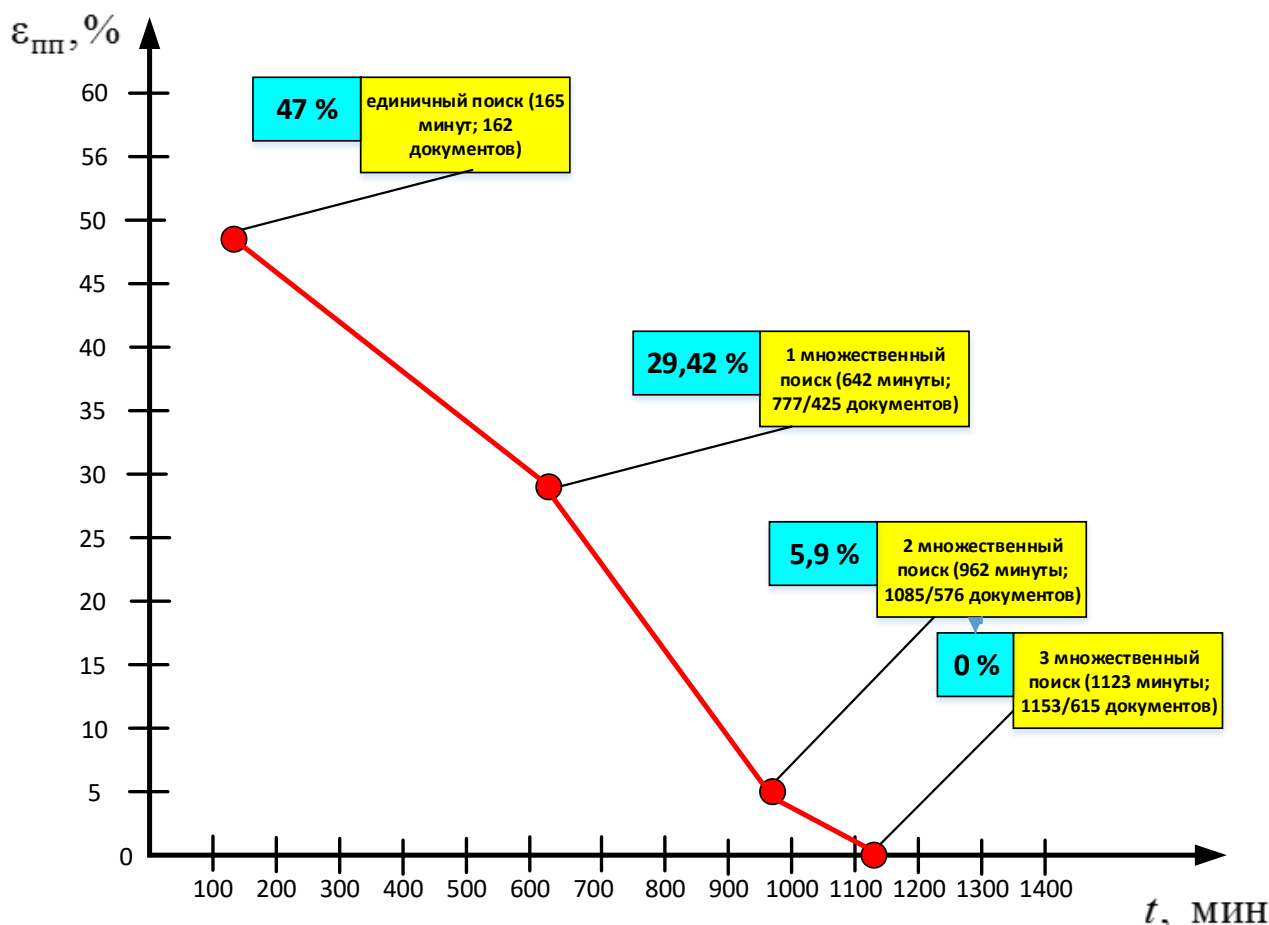


Рисунок 1. Зависимость величины ошибки патентного поиска от временных затрат

Выводы. Третий множественный поиск позволил свести ошибку патентного поиска от искажений и шумов патентного поиска к нулю. В тоже время величина ошибки относительно третьего множественного поиска увеличилась для: второго множественного поиска на 5,9 %; первого множественного поиска на 29,42 % и единичного поиска на 47 %.

Качество управления интеллектуальной собственностью напрямую зависит от достоверности и величины ошибки полученных результатов патентного поиска. Проведенный анализ показал, что величина ошибки патентного поиска зависит от технически грамотной замены слов при поисковом запросе относительно заданного технического результата и наличия временных ресурсов.

Литература.

1. Д.А. Журавлев, С.В. Мурашова, А.И. Литвинов Построение патентных ландшафтов как инструмент принятия управленческих решений в сфере научных исследований и разработок / Научный журнал «Человек и образование» – 2018. – № 5. С. 182-188.
2. Д. А. Журавлёв, С.В. Мурашова, А.И. Литвинов «Анализ проблем построения патентных ландшафтов в сфере научных исследований и разработок в ВС РФ и пути их преодоления». Инновационная деятельность в Вооруженных силах Российской Федерации: Труды всеармейской научно-практической конференции. 10-11 октября 2018 года, – СПб.: ВАС, 2018. 190 С. 24-28.

3. Отчет о научно-исследовательской работе «Исследование методик составления отчетов о патентных ландшафтах как инструмента принятия управленческих решений в сфере научных исследований и разработок» (шифр темы 9-ЭП-2014) / Е. В. Королева, А. А. Молчанова, Н. В. Повов, О. О. Вяль, М. В. Звягина, Н. И. Паршин, М. А. Якименко / ФГБУ «Федеральный институт промышленной собственности». – М.: ФИПС, 2016.
4. Л. Г. Кравец. Зарубежный опыт построения патентных ландшафтов. Научно-практический журнал Интеллектуальная собственность. Промышленная собственность, специальный выпуск. 2016, С. 96.
5. Д. А. Журавлёв, А. И. Литвинов, Н. С. Агеева. «Оценка величины ошибки от искажений и шумов патентного поиска при построении патентных ландшафтов». Инновационная деятельность в Вооруженных силах Российской Федерации: Труды всеармейской научно-практической конференции. 10-11 октября 2018 года, – СПб.: ВАС, 2018 – 190 С.17–21.
6. Bogdanova E., Maximova T.G., Nikolaev A.S., Antipov A.A. Innovation Management in the Organization Using Patent Analytics Tools in the Analysis of the Competitive Environment // Proceedings of the 31st International Business Information Management Association Conference (IBIMA). Innovation Management and Education Excellence through Vision 2020-2018, pp. 6856-6863. <https://ibima.org/accepted-paper/innovation-management-in-the-organization-using-patent-analytics-tools-in-the-analysis-of-the-competitive-environment/>