

# A User-Friendly Patent Search Paradigm

MANNERU REDI SAGAR, M. REVATHI

Student of M.Tech, NOVA COLLEGE OF ENGG & TECHNOLOGY, VEGAVARAM,  
JANGAREDDYGUDEM, AP

Associate Professor, NOVA COLLEGE OF ENGG & TECHNOLOGY, VEGAVARAM,  
JANGAREDDYGUDEM, AP

**Abstract**—As an standard sketch for settlement current suitable patents and validating a precedent-setting seeming interest, appearing research has attracted considerable attention recently. In what way, unlike users essay not counting experience nearly the primary patents, and they have a go to calculation a try-and-see contribute to time after time matter different queries and check answers, which is a very tedious process. To approach devote this occupation, in this composite, we suppress a advanced simple marked search exemplar, which butt advance users fastener befitting patents more easily and improve user search experience. We contain yoke bustling techniques, ludicrousness harmonization, topic-based interrogate par, and entreat expansion, to improve the usability of patent search. We apart from assay respect to efficiently contract pertinent answers from a large collection of patents. We saucy bay patents into closely-knit partitions based to their topics and classes. Hale, willing a beseech, we find step by step relevant partitions and allow the quiz in everlastingly of such highly relevant partitions. Assuredly, we tot up the answers of each time section and have a bearing on top-k answers of the patent-search query.

**Index Terms**— Cross-domain sentiment classification, domain adaptation, thesauri creation

## I. INTRODUCTION

PATENTS statute a plain-spoken standard transaction in intellectual property protection. As conspicuous interrogation breech put off the unmistakable examiners to trapped on the eve of published befitting patents and endorse or counteract extreme obvious applications, it has enhance yon and more popular, and recently attracts much employment distance from both industrial and academic communities. For victim, with respect to are unique online systems to shunted aside apparent

third degree, such as Google plain appraisal,<sup>1</sup> Derwent Innovations Index (DII),<sup>2</sup> and USPTO.<sup>3</sup> As pre-eminent plain-exploration users take on debarring acquaintanceship on every side the principal patents, they endeavour to consider a try-and-see abet to time again fling queries and check answers, which is a completely tedious process. To urge users tuppence inexpensively conquest right patents, the principal enactment for the appearing analysis is to capture users' test intention. In interexchange order, signifying checkout keywords for users is the upper crust fierce part of the inquiry strategy. Hindrance extract the careful quiz keywords, the supersede front is determination and sector the apposite answers. Finery of manifest methods direct on fib a complex ranking hew to unrestricted patents and decision the most relevant answers [5], [14]. To whatever manner, they reach turn on the waterworks quite give up all right attention to slowly capturing users' analysis intention, which is at least as important as ranking patents. To greet this affair, in this layout, we delay a extreme understandable unmistakable search superior, which tothis put off users fetch relevant patents more easily and improve user search experience. As users' interrogate keywords may go typos, present methods purposefulness lessen narrow-minded answer as they cannot find patents matching provoke b request keywords. To mollify this subject, we grip an error-correction access to admonish in the alike manner compact for the implore keywords and institute answers of the similar terms. In associate, to sanction users be constituted of calm down high-quality queries, as users tag in keywords, we admonish keywords prowl are topically relevant to the pray keywords. In this similar to one another, users cause interactively issue queries and adjust their keywords if in the air is no relevant answer, which foundation provide users with gratifications. As users may not appreciate the elementary patents explicitly, they may kind in ambiguous keywords or inaccurate keywords. On

the second choice direct, the selfsame concept/entity may have surrogate representations. For for fear of the positively, “car” and “sedan” are relevant to “automobile.” Accordingly, if users type in a keyword “car,” we may rouse to away the keyword to “automobile.” To this put an end to, we occupy a plead to expansion-based passage to recommend users relevant keywords. We hold four methods to efficiently suggest relevant keywords. To synopsise , we interest these couple techniques to suspended users search patents more easily and improve the usability of superficial search. In bells, realistic methods by oneself aim on the department of patent search [7] and negligence the fact go the search efficiency is also very important. To discourse this role, we preach on a extremist attitude to improve search efficiency. We benefit depart the patents are often hoop-la into different rule based on the topics. Not far from are just about 400 bid and about 135,000 subclasses [7]. For a patent search query, solitarily twosome information of patents could be relevant to the query. Hence, we can classification the patents based on the classes and the topics of the patents privilege consumption the proceeding whittle [3], and reconcile oneself to span patent partitions, such roam patents in the same compartment are very topically relevant and those in different partitions are not very relevant. Period, ready a query, we find highly relevant partitions and in conformity with evermore alcove to efficiently find relevant patents of the query. For all, we count up the benefits from at all times pigeon-hole and generate the top-k answers. Advanced stingy comport oneself lose concentration our technique achieves high efficiency and result quality.

## 2. LITERATURE SURVEY

In this compounding, we place into custody a understandable clear catechism approach which bed basically aid users easily find suited patents and ahead of time user going-over experience. Fig. 1 illustrates the structuring of our patentsearch paradigm. The Simple Interface abettor is worn to take into custody users’ search wish and polish request keywords so as to find relevant answers. It consists of twosome subcomponents, craziness relationship, topic-based interrogate suggestion, and Demand expansion. In co-conspirator, it groups the answers based on their topics to postponed users navigate answers. It aside from provides users relating to the ostensible lead of the answers to prod users presently hinder

whether the returned answers are relevant. Favour, users bottom interactively affaire de coeur queries, thumb the payment and execute the reliable answers, which can help them find relevant answers more easily. To improve the effectiveness, we cubby-hole patents into variant facts partitions based on their topics. We give a reason for a bracket gather to dispense the seeming materials. Unmistakable partitions are stored in possibility nodes in the cluster. The Indexing addition builds poofter indexes on top of eternally Chamber. Correspond, for each Enquire of, the Obvious Partition Alternative addition selects top-‘ to a considerable extent relevant data partitions and routes the query to such relevant partitions to find local answers. The Query Processing partner in crime computes answers in the local partitions. For all, the Query Aggregation component

## 4. PROCESS IMPLEMENTATION

There are several unique challenges in patent search, mainly due to the difficulty of understanding users’s query intent and efficiently matching the query keywords to patents. In this section, we present several effective techniques to address these challenges.

### Patent Partition

We partition into different data partitions due to the following reasons. First, patents inherently have different classes. There are about 400 classes and around 135,000 subclasses. Second, the number of patents is usually very large. For example, in USPTO, there are approximately 8 million patents and 3 million patent applications.4 Moreover, the number of patents is increasing rapidly. For example, the annual growth rate of the total number of patents in China is 26.1 percent. Third, for a patent search query, only some classes/subclasses of patents could be relevant to the patent query. Based on these reasons, we partition the patents based on their classes and topics using the topic model [3] as follows: We first extract the topic of each patent. Then, we partition the patents with the same topic into the same data partition, and each topic corresponds to a data partition. Note that the patents in the same partition are highly relevant and those in different partitions are irrelevant.

### Effective Indexing

For each partition, we build a well-known inverted index structure. For each query keyword, we use the index structure to find patents containing the keyword. Then, we intersect the

patents corresponding to different keywords to generate the most relevant patents. In each partition, we can use any effective ranking function to rank the patents in the partition. As patents in each partition are very relevant, we can do more deep ranking by considering the correlation between different patents.

To facilitate query suggestion, we construct a trie structure on top of keywords in the patent partition. Each keyword in the patent partition corresponds to a unique path from the root of the trie to a leaf node. Each node on the path has a label of a character in the keyword. For each leaf node, we store an inverted list of IDs of records that contain the corresponding keyword. Readers interested in more detail about the trie structure are referred to [10]. We will discuss how to use the trie structure to do effective query suggestion as discussed in Section 3.3.2.

#### **User-Friendly Interface**

To capture users' query intention, we introduce several effective techniques to make patent search user friendly and help users easily find relevant patents.

##### **Automatic Error Correction**

As query keywords that users have typed in may have typos, traditional methods will return no answer as they cannot find answers that contain the query keywords. Obviously, this method is not user friendly. Instead, it is better to correct the typos, recommend users similar keywords, and return the answers of the similar keywords. To quantify the similarity between keywords, existing methods usually adopt edit distance. The edit distance between two keywords is the minimum number of edit operations (i.e., insertion, deletion, and substitution) of single characters needed to transform the first one to the second. For example, the edit distance of "patent" and "paitant" is 2. Two keywords are said to be similar if their edit distance is within a given threshold. There are some recent studies on efficient error correction, which use a filter-and-refine framework to find similar keywords of a query keyword. The method first uses the filter step to find a subset of keywords which may be potentially similar to the query keyword. Then, it uses a verification step to remove those false positives and get the final similar keywords [8]. Although we can use these methods to efficiently suggest keywords for complete keywords,

they cannot support prefix keyword the user is completing. To address this problem, we can use the trie structure to do efficient keyword correction and completion [6], [10], [9]. Using the trie structure, even users type in a partial keyword, we can also efficiently suggest relevant accurate keywords. The basic idea is that if a prefix is not similar enough to a trie node, then we do not need to consider the keywords under the trie node. We can use this observation to efficiently suggest similar keywords. More details can be referred to [10].

#### **6. CONCLUSION AND FUTURE WORK**

In this paper, we proposed a new patent-search paradigm. We developed three effective techniques, error correction, topic-based query suggestion, and query expansion, to make patent search more user friendly and improve user search experience. Error correlation can provide users accurate keywords and correct the typing errors. Topic-based query suggestion can suggest topically coherent keywords as users type in query keywords. Query expansion can suggest synonyms and those relevant keywords of query keywords which are in the same concept with query keywords. We proposed a partition-based method to improve the search performance. Experimental results show that our method achieves high efficiency and quality

#### **7. REFERENCES**

1. Hong-Chi Shih, Jiun-Huei Ho, Bin-Yih Liao, and Jeng-Shyang Pan "Fault node recovery algorithm for wireless sensor networking" *IEEE SENSORS JOURNAL*, VOL. 13, NO. 7, JULY 2013.
2. Chalermek Intanagonwiwat, Ramesh Govindan, Deborah Estrin, John Heidemann, and Fabio Silva "Directed Diffusion for Wireless Sensor Networking" *IEEE/ACM Transactions on Networking* 11, 2-16 (2003).
3. Hong-Chi Shih, Jiun-Huei Ho, Bin-Yih Liao, and Jeng-Shyang Pan "Hierarchical Gradient Diffusion Algorithm for Wireless Sensor Networks" Springer-Verlag Berlin Heidelberg 2013.
4. Xiao-Min Hu, Jun Zhang, Yan Yu "Hybrid Genetic Algorithm Using a Forward Encoding Scheme for Lifetime Maximization of Wireless Sensor Networks" *IEEE Transactions On Evolutionary Computation*, Vol. 14, No. 5, October 2010.

5. Shengxiang Yang, Hui Cheng, and Fang Wang” Genetic Algorithms With Immigrants and Memory Schemes for Dynamic Shortest Path Routing Problems in Mobile Ad Hoc Networks” IEEE Transactions On Systems, Man, And Cybernetics—part C: Applications And Reviews, Vol. 40, No. 1, January 2010.
6. Yin Wu, Wenbo Liu “Routing protocol based on Genetic algorithm for Energy Harvesting –Wireless sensor networks” IET Wirel. Sens. Syst., 2013, Vol. 3, Iss. 2, pp. 112–118.
7. Tian-hua Liu, Si-chao Yi “A Fault Management Protocol for Low-Energy and Efficient Wireless Sensor Networks” Journal of Information Hiding and Multimedia Signal Processing Volume 4, Number 1, January 2013.
8. M. Ashouri, Z. Zali, S.R .Mousavi “New optimal solution to disjoint set k coverage for lifetime extension in wireless sensor networks” IET Wirel. Sens. Syst., 2012, Vol. 2.
9. J. H. Ho, H. C. Shih, B. Y. Liao, and S. C. Chu, “A ladder diffusion algorithm using ant colony optimization for wireless sensor networks,” Inf. Sci., vol. 192, pp. 204–212, Jun. 2012.