Similarity Analysis of Legal Judgments

Sushanta Kumar
International Institute of Information Technology, Hyderabad, India
sushanta.kumar@research.iiit.ac.in

P. Krishna Reddy
International Institute of Information Technology, Hyderabad, India
pkreddy@iiit.ac.in

V. Balakista Reddy and Aditya Singh
NALSAR University of Law, Hyderabad, India
balakista@gmail.com

ABSTRACT

In this paper, we have made an effort to propose approaches to find similar legal judgements by extending the popular techniques used in information retrieval and search engines. Legal judgements are complex in nature and refer other judgements. We have analyzed all-term, legal-term, co-citation and bibliographic coupling-based similarity methods to find similar judgements. The experimental results show that the legal-term cosine similarity method performs better than all-term cosine similarity method. Also, the results show that bibliographic coupling similarity method improves the performance over co-citation approach.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing - linguistic processing; H.2.8 [**Database Management**]: Database applications - Data mining; J.1 [**Administrative Data Processing**]: Law

General Terms

Algorithm, Experimentation

Keywords

Information retrieval, Case retrieval, Legal judgments, Case citation

1. INTRODUCTION

Automatic information or document retrieval systems are the main components of modern decision support systems or search engines to reduce the information overload. Investigating the approaches for improving the performance of document retrieval systems and search engines is an active research area [3].

In this paper we have made an effort to investigate the approaches to find similar legal judgments under common law system [2]. The objective is to build a legal decision support system to help the stakeholders such as judges and lawyers to quickly find the similar judgments to improve the efficiency. Typically, a lawyer needs to analyze older judgments similar (or relevant) to the current case to prepare the arguments. In order to find similar judgments, a lawyer searches the database of judgments.

In the area of legal information systems, research efforts are going on to investigate effective summarization methods [10] [7]. Recently an approach [11] has been proposed in which case citations of the judgments are exploited to form semantic based network which allows legal professionals to find the related documents.

Note that a judgment is of fairly big size, complex and contains domain specific format and semantics. In addition, it also refers to other judgments in case of common law system. Developing a method to compare two judgments is an important issue. Traditional information retrieval methods compare two documents by treating them as bag-of-words. In these methods each term of a document is given weight according to tf-idf method [8]. The vector space model and its variations [9][5] are used to find the similarity. Also, links are being exploited to compare two web pages using the notions of co-citation [4] and bibliographic coupling [6] methods. Similarity citations are exploited in case of research papers.

In this paper we have extended four types of similarity methods related to information retrieval and web environment and analyzed the performance of these methods in extracting similar legal judgments. The experiments were conducted on the judgment dataset of Supreme court of India. The user study results show that, both legal term and bibliographic based similarity methods exhibit better performance.

The rest of the paper is organized as follows. In the next section we explain about legal judgments. In section 3, we present similarity measures. The experimental results are discussed in section 4. The last section contains conclusions.

2. ABOUT LEGAL JUDGMENTS

Some of the significant parts of a judgment are: 'Act', 'Headnote', 'citation' and 'case citation'. The term 'Act' indicates the category of the dispute discussed in the judgment. For example, a judgment with 'Land acquisition Act' indicates the dispute about land. The 'Headnote' part consists of a brief summary of judgment. The term 'Citation' refers to the unique IDs by which a judgment will be referred by other judgments. For example consider the following citation: "(1984) 1 SCC 339". In this citation '1984' is the year

COMPUTE 2011 March 25-26, Bangalore, Karnataka, India

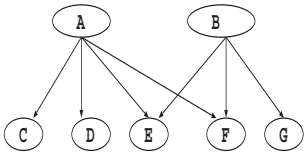

Figure 1: An illustration of case-citations of two Judgments.

of publication, '1' is the volume number of reporter, 'SCC' (Supreme Court Cases) is the name of the law reporter and '339' is page number of the judgment of the volume. The 'Case citation' indicates the references made to other judgments.

Significance of 'case citation': Typically 'case citations' contribute towards the argument of the judgments by leveraging the legal concepts of the cited judgments. Hence, all the cited judgments go through strict scrutiny of legal experts, during argument between two opposing lawyers.

For a given judgment there are two types of 'case citations', which are explained below.

Out-citation of a Judgment For a given judgment 'J', we define out-citations as those 'case citations' which are mentioned in judgment 'J' and are referring other judgments. In Figure 1, judgment 'A' has four out-citations namely judgments 'C','D','E' and 'F' while judgment 'B' has three out-citations namely 'E','F' and 'G'.

In-citation of a Judgment For a given judgment 'J', we define in-citations as those 'case citations' which are referring to judgment 'J'. In Figure 1, judgment 'A' is in-citation for judgment 'C','D', while judgments 'B' is in-citation for judgment 'G'. Judgments 'A' and 'B' both are in-citations for judgments 'E' and 'F'.

3. SIMILARITY MEASURES

The four similarity measures employed in this paper to compare the judgments are as follows.

(i) All-term cosine similarity: All terms in the judgments are extracted by following information retrieval approaches based on tf-idf score. A judgment is considered as a vector of terms. Vector space model [9] is used for similarity computation.

(ii) Legal-term cosine similarity: In a given judgment, only those terms that appear in legal dictionary are considered as representative terms. For each legal term tf-idf score is computed. A judgment is considered as vector of legal terms. Vector space model [9] is used for similarity computation.

(iii) Bibliographic coupling similarity: Out-citations are identified using regular expression for each law reporter and extracted by scanning the headnote of each judgment. For each judgment, <judgment, out-citation> pair is prepared. Bibliographic coupling similarity score between two judgment is equal to the number of common 'out-citations'.

(iv) Co-citation similarity: After extracting out-citations, as mentioned above paragraph, <judgment, out-citation> pair is converted to <judgment, in-citation>. Co-citation similarity score between two judgments is equal to the number of common 'in-citations'.

4. EXPERIMENTS

Experiments were conducted on a dataset of 2,430 judgments delivered by Supreme court of India (Common law system)[1]. Number of out-citations of judgments in our dataset are between 3 to 12.

We performed user study to evaluate the performance. Sample pairs (of judgments) were given to each one of the 5 legal domain experts without informing the computed similarity values. Domain experts were asked to give the similarity score between 1(low) to 10(high). Analysis was done after averaging the score given by experts. Legal domain experts assigned similarity score of judgment pairs based on three criteria namely, similarity in issue discussed in the judgment, similarity in underlying facts of the judgment, and utility to the lawyer researching for judgments similar to a given judgment.

First experiment: We conducted experiments to analyze the performance of four similarity measures. In this experiment, we verified the relative significance of each similarity measure. We selected four different categories of sample pairs of judgments such that each category contains high similarity score using one feature while lower similarity score using remaining features: first one is with high cosine similarity score based on 'all terms' feature vector, second one is with high cosine similarity based on 'legal term' feature vector, third one is with co-citation similarity score equal to 3 and the fourth one is with bibliographic coupling similarity score equal to 3.

The computed similarity values are compared with the similarity values (normalized to 0 to 1) given by domain experts. The results of first experiment is shown in Table 1.

From the results on sample I, it is evident that the domain experts are not agreeing to the computed similarity values based all-terms feature vectors. It is because, owing to its verbose nature, a judgment contains large number of texts to explain the applied legal concept in full detail. Since the a legal concept can be explained using various combination of words, text based similarity methods fail to satisfy human notion of similarity.

The sample II results show that, the domain experts agree with computed similarity values much better than all-terms case. High similarity score by domain experts indicate that legal-terms are capturing the essence of judgment more effectively than all-terms vector. Typically, if available, a judge employs legal terms from legal domain for expressing the applied legal concepts to improve communication and understanding. So it is natural that, any similar judgments will have common legal term. As a result, similarity computation based on legal-terms are giving better results over all-terms.

Analysis of sample III results shows the results based on bibliographic coupling similarity score. Results show that domain experts agree with the values computed based on bibliographic coupling similarity. Our observation shows that, if two judgments cite the same judgments, both agree to the context of the judgment. A typical judgment is popular for certain subtopics. So, if two judgments cite the same judgment, they also agree on the subtopic most of the cases.

Co-citation similarity performance can be analyzed from sample IV results. It is noted that the domain experts are not agreeing with the computed similarity values based on co-citation similarity. This is a surprising result as compared to the phenomena in scientific literature. Since, a judgment is divided into several subtopics to cover the diversity of the

issue and a reference will be made for each subtopic. If a judgment is citing two judgments, it does not mean that two judgments are similar as judgment may cite those two judgments for two different subtopics altogether.

Second experiment: We explored the effectiveness of bibliographic coupling method with another experiment. Since judgments in the dataset contained minimum 3 out-citations, so the maximum value of out-citation score which will include all the judgments from our dataset will be 3. So all the judgment pairs having bibliographic coupling score $\geq$ 3 in our dataset, were given to legal domain experts. The results are shown in Table 2. It can be observed that 17 out of 18 pairs matches with the similarity score (i.e ≥ 0.50) given by domain expert. The percentage of accuracy comes to 94.44%.

The results indicate that the legal-term similarity score and bibliographic coupling score satisfy the human notion of similarity. Since being a domain specific text document, applicability of legal-term cosine similarity score is quite expected one, high accuracy of bibliographic coupling score comes out as an interesting result.

Table 1: Performance of similarity measures.

Sample type	Judgment pairs IDs	All terms score	Legal term score	BC score	CC score	Expert score
I	1	**0.95**	0.60	0	0	0.20
	2	**0.98**	0.07	0	0	0.40
	3	**0.89**	0.15	1	0	0.40
	4	**0.81**	0.17	0	0	0.20
	5	**0.86**	0.35	0	0	0.30
II	6	0.15	**0.98**	2	0	0.80
	7	0.07	**0.99**	1	2	0.90
	8	0.43	**0.95**	0	2	0.80
	9	0.47	**0.97**	0	0	0.90
	10	0.33	**0.73**	0	1	0.60
III	11	0.22	0.31	**3**	0	0.80
	12	0.27	0.33	**3**	0	0.70
	13	0.42	0.53	**3**	0	0.80
	14	0.22	0.48	**3**	0	0.60
	15	0.07	0.24	**3**	0	0.50
IV	16	0.11	0.14	0	**3**	0.30
	17	0.21	0.35	0	**3**	0.40
	18	0.13	0.17	0	**3**	0.30
	19	0.24	0.37	0	**3**	0.40
	20	0.34	0.44	0	**3**	0.30

Analysis by domain experts: Here, we present the views of the domain experts on three judgment pairs of Table-2 regarding justification for the given similarity score.

Pair 1989_164 & 1990_24: The pair received a low similarity score by domain expert and average text based similarity score. The justification by domain expert is as follows.

In Case 1989_164, the Court is looking at issues relating to the principle of 'res judicata'. There is a discussion on the application of the principle in cases where petitioners try to challenge the validity of the provisions of the Act on different grounds at different times. In Case 1990_24, the judgment revolves around issues of pending suits for eviction. The reason for a low domain expert score is basically that there is very little substantive similarity. The facts of the two cases are distinct. The facts, issues and legal principles are not closely connected. The reason why a search engine may throw up these two cases as being similar is because of similar legislations or the use of same terms ("land", "rent" etc) in the two judgments.

1989_122 & 1990_56: This pair received average domain expert similarity score and a very low text based similarity score. The justification is as follows.

While Case 1989_122 deals with the issue of discrimination in promotion and pay scales on the basis of educational qualification. The central issue here was whether two groups of employees, one that includes degree holders and the other with diploma holders, be treated equally in promotion and payment of salaries.

In Case 1990_56, the Court was dealing with petitions claiming the application of "equal pay for equal work" principle for those employees who had not been regularized but were in service for a long period of time.

These cases are similar in that they involve a close discussion on the legal principle of "equal pay for equal work". The Supreme Court in both these cases considers the constitutional scheme of the directive principles of state policy while discussing the principle. However, there is not a lot of similarity in the facts and the judgments also look at distinct issues - the first case has a discussion on the nature of the directive principles while the second case looks at issues of taxation. These cases will be fairly useful for a researcher or a lawyer who wishes to argue or cite cases on the principle of equal pay for equal work but not so much in other issues. Hence, the average similarity score is given.

Pair 1989_83 & 1990_56: This pair has received reasonably high similarity score by domain expert and a very low text based similarity score.

In Case 1989_83, the Court was addressing a petition by a hearing therapist, who claimed that while he was performing a job same or similar to senior speech pathologist, senior physiotherapist, senior audiologist, and speech pathologist in the same institution under the same employers, he had been given a lower pay scale in comparison to these posts. The Court adjudicated the reliance on the equal pay for "equal work principle".

In Case 1990_56, the Court was dealing with petitions claiming the application of "equal pay for equal work" principle for those employees who had not been regularized but were in service for a long period of time. The Court here agreed with the petitioners but cautiously gave directives to the State Government keeping in view the economic capacity of the States.

For most lawyers and legal researchers, these two cases are going to be useful because the core principle is the same. There is an extensive discussion of the various arguments for and against the application of the principle of "equal pay for equal work" and the Court looks at the elements of this principle in-depth.

Summary of analysis: The analysis indicates that bibliographic coupling similarity score is capturing the similarity notion used by domain experts.

5. SUMMARY AND CONCLUSION

In this paper, we have analyzed the issue of finding similar legal judgments by extending the similarity measured used in information retrieval and web documents. The results

Table 2: Performance of bibliographic coupling similarity $\geq$ 3.

Sl. No.	Judgment pairs	All terms score	Legal term score	BC score score	CC score	Expert score
1	1986_232 & 1992_196	0.07	0.24	3	0	0.50
2	1987_113 & 1989_2	0.04	0.04	3	1	0.50
3	1987_202 & 1990_31	0.16	0.12	3	0	0.70
4	1989_122 & 1990_56	0.07	0.08	3	2	0.50
5	1989_164 & 1990_24	0.28	0.18	3	0	0.40
6	1989_249 & 1990_31	0.18	0.73	3	0	0.50
7	1989_367 & 1992_161	0.53	0.84	3	0	0.60
8	1989_385 & 1990_31	0.12	0.51	3	0	0.70
9	1989_75 & 1990_249	0.23	0.32	3	1	0.80
10	1989_75 & 1990_284	0.36	0.18	3	1	0.80
11	1989_83 & 1989_122	0.42	0.53	3	0	0.80
12	1989_83 & 1990_56	0.06	0.07	3	2	0.70
13	1990_31 & 1990_122	0.42	0.67	3	0	0.80
14	1990_31 & 1990_284	0.27	0.33	3	0	0.70
15	1990_284 & 1993_206	0.22	0.48	3	0	0.60
16	1990_82 & 1991_55	0.60	0.79	3	0	0.60
17	1993_236 & 1993_241	0.99	0.95	3	0	0.60
18	1989_75 & 1990_31	0.22	0.31	3	0	0.80

show that all-term cosine method and co-citation similarity methods are not efficient in finding similar judgments. The legal-terms cosine and bibliographic coupling similarity methods are efficient in finding similar legal judgments. As a part of future work, we are planning investigate a scalable approach to find similar legal judgments.

6. REFERENCES

[1] Supreme court of india judgments. http://www.commonlii.org/in/cases/INSC.

[2] Wiki page, 2010. http://en.wikipedia.org/wiki/Common_law.

[3] W. B. Croft. What do people want from information retrieval? Technical report, 1995.

[4] S. H. Co-citation in the scientific literature: A new measure of the relationship between two documents. In *Journal of the American society for Information Science*, pages 24:265–269, 1973.

[5] K. Hammouda and M. Kamel. Phrase-based document similarity based on an index graph model. In *Data Mining, 2002. ICDM 2002. Proceedings. 2002 IEEE International Conference on*, pages 203 – 210, 2002.

[6] M. Kessler. bibliographic coupling between scientific paper. In *American Documentation*, pages 14(1):10–25, 1963A.

[7] M.-F. Moens. Summarizing court decisions. *Inf. Process. Manage.*, 43(6):1748–1764, 2007.

[8] G. Salton and C. Buckley. Term weighting approaches in automatic text retrieval. Technical report, 1987.

[9] G. Salton, A. Wong, and C. S. Yang. A vector space model for automatic indexing. *Commun. ACM*, 18(11):613–620, 1975.

[10] M. Saravanan, B. Ravindran, and S. Raman. Improving legal information retrieval using an ontological framework. *Artif. Intell. Law*, 17(2):101–124, 2009.

[11] P. Zhang and L. Koppaka. Semantics-based legal citation network. In *ICAIL '07: Proceedings of the 11th international conference on Artificial intelligence and law*, pages 123–130, 2007.