|| Volume 3 || Issue 2 || February 2018 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

Trademark Approval System Using Semantic Similarity Computation

Mr. S. S. Sanghavi[1] ,Asst.Prof. V.S. Karwande[2]
Student ME-II, Department of Computer Science and Engg, *DSQIET, Aurangabad, India[1]*
Asst Prof., Department of Computer Science and Engg. *DSQIET, Aurangabad, India[2]*
vjy725@gmail.com[1], sagarssanghavi@gmail.com[2]

Abstract **Trademarks are signs of very high reputational value as asset. Thus, they require some protection. A trademark is a mark that one can use to recognize business products and services from other vendors. The rapid growth in e-commerce at early in 21st century has had a remarkable impression on intellectual property and management. A particular area of concern is the misuse of trademarks and trademark requires protection. Trademarks are possessory words with high reputation. The paper presents the review of different techniques used for similarity measures and their contrasting factors of Trademarks.**

Keywords: *- conceptual similarities, trademark comparison, Semantic Analysis, feature extraction***.**

I INTRODUCTION

Trademarks, as defined by the European Office of Harmonization in the Internal Market (OHIM), are signs that are used in trade to identify products or services. They have become intangible intellectual property (IP) assets that allow goods or services to be easily recognized by consumers. The number of trademarks registered and used each year in the marketplace shows an upward trend with no significant sign of declining. Trademark infringement is a form of IP crime that may lead to serious economic problems. In general, IP-intensive companies make twice as many sales as non-IP-intensive companies. Searching for conceptually similar trademarks is a text retrieval problem, and then traditional text retrieval systems based on keywords are not able to retrieving conceptually related text. This limitation motivates research in the semantic technology. Few common outcomes from trademarks infringement is lost income, low benefits, and need extra money of conservancy to stave off next infringement. The trademarks registered improve by 20 percent from last many years in the word. Trademark similarity problems for the other 70 percent stay deficiently researched in more that content-based retrieval goes from different limitations.

When assessing trademark infringement cases then analysis several separate components, such as the same of the goods, the especial and main points of the different trademarks, and the similarity of the trademarks. A trademark may be designated by the following symbols: is trademark symbol, which is the letters TM, for an unregistered trademark, a mark used to promote or brand goods is the letter R surrounded by a circle, for a registered trademark. Infringement may occur when the infringer, uses a trademark which is confusingly similar to a trademark owned by another party. These searches look for trademark that matches some or all words in a question line wording. As indicated in their latest printing on trademark knowledge-bases and look for systems. Two trademarks are necessary not same to make an infringement. The conceptual different of text files that part of same domain, utilization same notations, or demonstration same consideration has been used broadly.

Some major damage resulting from trademark infringement is lost revenue, lower profits, and the additional cost of protection to avoid future infringement. In a statistics provided by the U.S. International Trade Commission, as reported by the Chairman of the Joint Economic Committee, the number of investigated infringement cases rose by 23.2% from 2010 to 2011. In 2012, a total of 3400 trademark infringement cases were filed in the U.S. District Courts. This does not include the presumably larger number of cases in which settlements are reached prior to the filing of cases [4]. In the same year, the European Commission also reported that trademark infringement accounted for the majority of IP crime, comprising about 97% of IP crime cases that year [5]. A compulsory analysis required by both European law and U.S. legal practice, when assessing trademark infringement cases is the "likelihood of consumer confusion" analysis. The analysis is an overall assessment that involves several interdependent factors, such as the similarity of the goods, the distinctive and dominant elements of the conflicting trademarks, and the similarity of the trademarks. The similarity of the trademarks is assessed based on the visual, conceptual, and phonetic aspects of the conflicting trademarks. Trademarks that are similar enough in these respects to be confusing for the average consumer are more likely to cause infringement.

II LITERATURE SURVEY

The concept of similarity has become understood in trademark infringement cases. It is one of the most important analytical factors in such cases because it is in the similarity

between trademarks that the roots of the confusion normally lie. Moreover, similarity, in the context of trademarks, is also not binary but a matter of degree. The rule of thumb is that the higher similarity between the trademarks, then they will cause confusion. This paper addresses one of the aspects of similarity assessed during trademark analysis, which is conceptual similarity. The confusion in trademarks is based on the visual, phonetic or conceptual similarity of the marks. Issue is concerned, on the overall impression given by the marks, bearing in mind, inter alia, their distinctive and dominant components.

The underlying technology embedded in existing trademark search systems is primarily based on text-based retrieval. Such systems search for trademarks that match some or all words in a string text query. In a recently launched search system, the OHIM provides an option that allows users to search for trademarks in different languages [6].

The most common retrieval method employed in the existing trademark search system, as well as in many other multimedia search systems, is known as the keyword-based search. This search generally looks for keywords that have been tagged as predefined metadata among items in a database; it then returns words with similar matches. In text retrieval, text mining is performed for document classification, as well as for acquiring potentially useful knowledge from documents. Simple search tasks may work well with traditional information systems. The emergence of semantic retrieval technology was inspired by the limitations of traditional keyword-based retrieval. Semantic retrieval employs external knowledge sources, such as ontologies, to overcome the limitations of keyword-based systems.

This paper addresses one of the aspects of similarity assessed during trademark analysis, which is conceptual similarity. The confusion in trademarks is based on 1) text-based retrieval**,** 2) the visual, 3) phonetic and 4) conceptual similarity of the marks. Issue is concerned, on the overall impression given by the marks, bearing in mind, inter alias, their distinctive and dominant components.

1) Text-based retrieval

Such systems search for trademarks that match some or all words in a string text query. A recent system for counting short-text and sentence semantic similarity. The method is depends on the concept that the sense of a statement is create of nope mere the sense of its particular words, but also the anatomical path the words are concatenated. Thus hold on and connects syntactic concatenated. Thus hold on and connects syntactic and semantic data to count the semantic similarity of two phrases. Semantic data is given from the lexical resources. Syntactic data is get from a strong parsing procedure that searches the sentences in every phrase. A syntax based providence to calculate the semantic similarity between phrases or short texts. The concept on which the system is based on the sense of phrases is creating of nope mere the senses of its particular words, but as well the different words are concatenated. A method and a model for detracting and listing information from main language data. The main domain prototype depends on a hypothetic scale that is of a domain ontology, which define the domain information, and a lexical scale based on WordNet, that defines the domain glossary. The semantic data retrieval engine that created justification easy keyword-based problems, as well as natural language-based problems. The engine is also ability to develop the domain information, searching recent and same facts added to domain model. The in duration probe suggests that the method is efficient to many forms and define nations with accurate purity.

The existing trademark search systems is based on text-based retrieval. Such systems search for trademarks that match some or all words in a string text query. The advance search system that gives three types of search: word prefix, full phrase, and exact match. The most common retrieval method employed in the existing trademark search system, as well as in many other multimedia search systems, is known as the keyword-based search. The need of semantic retrieval technology was inspired by the limitations of traditional keyword-based retrieval. Semantic retrieval employs external knowledge sources, such as ontologies, to overcome the limitations of keyword-based systems.

2) Visual Similarity

F. M. Anuar, R. Setchi, and Y. K. Lai[1], author proposed Trademark image retrieval using an integrated shape descriptor. as the proposing innovatory trademark reflow technique to use the reform performance of expositor. The Zernike moment edge gradient technique (ZMEG) technique is used and in that used employed shape features and descriptor matching stage.

Visual similarity is focuses on sequence of the letters in that trademarks and font style variations. This example shows that these both signs are visually similar. Innovatory trademark reflow technique to use the reform performance of expositor. Trademarks are distinctive visual symbols with high reputational value, due to the perception of quality and innovation associated with them. They are important reputational assets used as a marketing tool to convey a certain assurance of quality, innovation, and the standards, which the manufacturer seeks to maintain. This motivates the need for trademark protection by providing a solution to prevent infringement. This problem can be addressed by developing retrieval systems capable of comparing the visual similarity of trademarks.

The visual similarity is checked during the trademark registration process where one of the steps involved is making sure that the trademark to be registered is not similar to any

trademark which is already registered. This is important in order to avoid infringements as well as to protect the rights of the existing trademarks. Author proposed to introduced substance point of a exclusive figure and this the point used to search nook pixel from it. Number of image collections available has increased due to ease of capturing images by different acquisition systems. The storage format of image data is relatively standardized however the effective retrieval of images from such databases remains a significant challenges. For the performance evaluation of the system we use the most commonly used method namely precision-recall. From the experimental result we conclude that the Trademark Image Retrieval based on shape feature perform better and gives satisfactory result.

3) Phonetic similarity

Phonetic aspect of similarity is common rhythm and intonation of the trademarks which considers the sound pattern and pitch variations in the syllable that form the trade marks. an algorithm to retrieve phonetically similar trademarks that can be used as a means for supporting trademark examination during the registration process. The algorithm employs a phonology based string similarity algorithm together with a typography mapping and token rearrangement to compute a phonetic similarity between trademarks. The trademark phonetic similarity score is then computed from the employed phonetic similarity algorithm. The proposed algorithm advances the state-of-the-art in trademark retrieval by providing a mechanism to compare trademarks with special characters or symbols phonetically. In general, there are two types of string matching: orthographical and phonological. Orthographic string matching estimates the similarity between two strings based on the number of steps required to transform one into the other or the number of characters they have in common. A commonly used measure is the edit distance measure, also known as the Levenshtein distance, which considers substitution, insertion, and deletion operations when comparing two strings. Alternatively, phonological measures employ phonetic features for each letter in the string and incorporate them into the similarity computation. Phonetic similarity measures have been widely employed in the areas of historical linguistics and genealogy. In the former area of study, phonetic algorithms are utilized to find similarities between languages. Examples include the study of similarities between cognates (i.e. words from different languages that share the same linguistic origin and etymology). In genealogy, phonetic algorithms are applied to name-matching applications, which attempt to retrieve closely similar names despite spelling variations.

4) Conceptual similarity

L. Sbattella and R. Tedesco [3], to presents a fact and ideal for substance and listing information from main data. Use the conceptual level and lexical level for describes the main information. The proposed system is provided good precision compare to regular search engine that is a simple and well powerful system.

Conceptual similarity aspect focuses on the semantic content of the trademark such as the meaning of the trademarks. The concept of similarity has become understood in trademark infringement cases. It is one of the most important analytical factors in such cases because it is in the similarity between trademarks that the roots of the confusion normally lie. Moreover, similarity, in the context of trademarks, is also not binary but a matter of degree. The rule of thumb is that the higher similarity between the trademarks, then they will cause confusion.

Methods

i. Edge Based Measure

Semantic similarity depends on the path length and on the position of the concept in the taxonomy. It employs the concept of common subsumers (i.e., the ancestor concept that subsumes the two concepts). It is simple to implement. Two concept pairs of equal length will have the same similarity. Two concept pairs that share exactly the same least common subsume and are of equal length will have the same similarity.

ii. Information Content

It assumes that the similarity between the two concepts can be derived based on the specificity of the concepts. The more specific a concept is in the taxonomy, the richer the information content will be. The information content calculation is derived based on the probability of the occurrence of concepts in the taxonomy. Two pairs with similar lcs and cumulative IC may have the same similarity.

iii. Feature Based Measure

It is independent of taxonomy and the subsumers of the concepts. It assumes that each concept has specific features that can be employed to measure similarity. It is defined as the ‘glosses’ (i.e., the definitions of concepts as the features that represent the concepts). The computational complexity is very high.

The work was motivated by increasing of fraud cases best an data similarities, where information retrieval system do not handle this particular issue and trademark similarity. The target on similarities during trademarks, which becomes when more than two or more trademarks like equal or relevant semantic implant. The advantages and limitations of each data similarity of reflow algorithm are described. The system work, conceptual similarities among trademarks like equal or relevant semantic implant. The desire of a hypothetic model of retrieval trademark is depends on hypothetical similarity. The main model language processing technology, data paths and lexical resources to calculate hypothetic similarity between different trademarks. The system is stimulated for improving of fraud cases best on data processing similarities, where data retrieval

system does not manage this particular problems. The system reforms on all ready trademarks find system by legislation a implementing of rectification the find to hypothetic same trademarks. The system employs natural language processing techniques, knowledge sources and a lexical resource to compute conceptual similarity between trademarks. Also confirm that the comparison of trademarks in terms of conceptual similarity.

III CONCLUSION

The work presented in this paper was motivated by the realization that despite the large number of infringement cases based on conceptual similarity, traditional information retrieval systems do not handle this particular issue well. It is also motivated by the understanding that trademark similarity, one of the factors that contributes to the likelihood of confusion, may be linked to the semantics of trademarks, i.e., their lexical meanings.

This paper contributes to the semantic algorithm to compare trademarks in terms of conceptual similarity. Entirely new similarity comparison concept in the domain of trademark retrieval. It utilizes natural language processing techniques, together with an external knowledge source in the form of a lexical ontology. The evaluation using both information retrieval measures and human judgment shows a significant improvement because the algorithm provides better results than the traditional baseline technique. The algorithm is not limited to the use of a specific word measure. This advantage provides the flexibility to choose any word measure suitable for particular applications or requirements.

ACKNOWLEDGMENT

The authors would like to thank Head of the Department who guide me and project coordinator and Project guide for their valuable guidance and support which helped to finalize many of the ideas behind this paper.

REFERENCES

[1] Office for Harmonization in the Internal Market, Annual Report 2012.[Online]. Available: https://oami.europa.eu, accessed Dec. 10, 2013.

[2] L. Dodell. (2013). The Trademark Problem: Casualty Insurance's Dirty Little Secret. [Online]. Available: http://www.carriermanagement.com, accessed Dec. 2013

[3] C. D. Scott, "Trademark strategy in the Internet age: Customer hijacking and the doctrine of initial interest confusion," J. Retail., vol. 89, no. 2,pp. 176–189, Jun. 2013.

[4] C. D. Scott, "Trademark strategy in the Internet age: Customer hijacking and the doctrine of initial interest confusion," J. Retail., vol. 89, no. 2, pp. 176–189, Jun. 2013.

[5] European Commission. (2012). Report on EU Customs Enforcement of Intellectual Property Rights. [Online]. Available: http://ec.europa.eu, accessed Dec. 1, 2013.

[6] P. Eagle, "News on patent, trademark and design databases on the Internet," World Patent Inf., vol. 34, no. 2, pp. 164–165, 2012.

[7] Y. S. Kim and W. Y. Kim, "Content-based trademark retrieval system using a visually salient feature," Image Vis. Comput., vol. 16, nos. 12–13, pp. 931–939, 1998.

[8] J. Schietse, J. P. Eakins, and R. C. Veltkamp, "Practice and challenges in trademark image retrieval," in Proc. 6th ACM Int. Conf. Image Video Retrieval, Amsterdam, The Netherlands, 2007, pp. 518–524.

[9] B. Furlan, V. Batanovic, and B. Nikolic, "Semantic similarity of short texts in languages with a deficient natural language processing support," Decis. Support Syst., vol. 55, no. 3, pp. 710–719, 2013.

[10] J. Oliva, J. I. Serrano, M. D. del Castillo, and A. Iglesias, "SyMSS: A syntax-based measure for short-text semantic similarity," Data Knowl. Eng., vol. 70, no. 4, pp. 390–405, 2011.

[11] L. Sbattella and R. Tedesco, "A novel semantic information retrieval system based on a three-level domain model," J. Syst. Softw., vol. 86, no. 5, pp. 1426–1452, 2013.