Result: Trademark approval system using semantic similarity computation

[1]Vijay Karvnde, [2]Sagar S. Sanghavi

[1]Asst. Prof, [2]Student M.E.,*CSE , DSQIET, Aurangabad, India*

[1]vjy725@gmail.com, [2]sagarssanghavi@gmail.com

Abstract: The rapid growth of e-commerce at the beginning of 21st century has had an important effect on intellectual property rights and management. A specific area can be misuse of trademarks and trademark shield. Trademarks are registered words and images with high worthy value; these are key assets, mostly used as a marketing tool, which need breach protection. This paper studies conceptual similarities among trademarks, which arises when two or more trademarks suggest matching or equivalent semantic content. This paper concern about the state-of-the-art by recommending a computational method based on semantics that can be used to match trademarks for conceptual similarity. A trademark retrieval algorithm developed and applied. The proposed retrieval algorithm is authenticated using two resources: a trademark database disputed cases and a database company names. In specific, the conceptual similarity of trademarks is a zone not ever before studied in information retrieval. This paper efforts on this significant feature by proposing a conceptual model of the comparison process, intended at retrieving conceptually similar trademarks. The proposed model processing and semantic technology to calculate the conceptual similarity between trademarks.

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. For example, in 2012, the OHIM received about 108 000 trademark applications, an increase of 2% from the previous year [1]. In the U.S., about 1 867 353 trademarks were registered and maintained during the first quarter of 2013, as compared with a total of 1 752 599 registered and in-use trademarks in the first quarter of 2012 [2]. The newly registered trademark statistic in the U.S. climbed by 10% from the 2010 fiscal year to the 2012 fiscal year [2]. 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. 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[3]. Searching for conceptually similar trademarks is a text retrieval problem, 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 match similarity issues and problems for the other 70 percent keep on deficiently investigate in more that of 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 might be selected by the following signs and symbols: is trademark signs and logos symbol, TM which is the letters Trademark, for an unregistered trademark, a mark used to promote or brand goods is the letter R surrounded by a circle, for a registered trademark. Violation may arise when the infringer, uses a trademark that arise confusingly similar to a trademark owned by a different 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.

The Impact of Intellectual Property Theft on the Economy

- IP infringement harms to companies by lost revenue, the costs of IP protection, damage to brand value, and decreased incentives to innovate because of potential theft.
- Consumers are harmed when they purchase duplicate product of lower quality, some of which, such as duplicate medicines, may cause health or safety risks. Governments lost tax. IP infringement reduces economic growth, weaken the nation's competitiveness, and decrease job creation.

Concept of similarity

The concept of similarity has become understood in trademark infringement cases. It is individual of the most essential investigative factor in such case as it is in the similarity among trademarks that the extraction of the confusion usually lies. 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.

II. LITERATURE SURVEY

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 major area of prototype depend on a hypothetic and logical scale that is of a area or 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 data reflow technique utilizes keywords passed by the user as the find measurement to find documents. Nevertheless, the language used in files is mostly hard and unclear, and hereby the outcomes obtained by using keywords are mostly not good. The way of this issue, created a semantic-based content mapping mechanism for a data reflow technique. These views simplify the find process and improving the purity of the returned results. The problems define during infringement litigation is the visible, hypothetic and phonetic similarity of different trademarks. This is focuses on important fact by defining a hypothetic model of the comparison process, target when retrieving hypothetic similar trademarks. The proposed model normal language accessing and semantic technology to get the hypothetic similarity between two or more trademarks. structure recommends a hypothetic mold of trademark retrieval base on top of the hypothetic match and similarity measures . The projected model extend on previously trademark finding models by providing discover to hypothetically related trademarks.

III. PROPOSED SYSTEM

The proposed retrieval algorithm is based on a conceptual model of the trademark comparison process developed in [10]. This paper focuses and extends the theoretical model by emergent and calculates a semantic algorithm for trademark retrieval based on conceptual similarity. The proposed algorithm applies Neuro-Linguistic Programming (NLP) procedure along with the word similarity distance scheme, which was derived from the WordNet ontology, together with a new trademark comparison measure. WordNet engaged in this algorithm owing to its lexical associations, that reflect human semantic association, and since it has also been verified successful in numerous previously developed works. The trademark comparison measure is derived from the Tversky contrast model [11], a well-known model in theory of similarity.
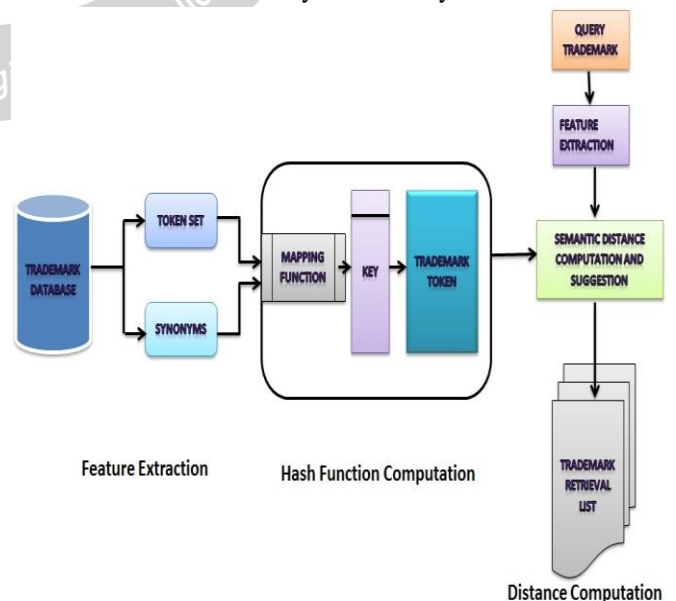

Figure 1 Proposed system model.

Step 1: Extracting Feature Extraction.

Every trademark signifies by two type of features (i.e., the trademark token, ft, and the synonym list, fs). The feature extraction step begins with a spelling correction process that corrects any spelling mistakes using a spellchecker. Then, frequent words (i.e., "no," "and," "the," etc.) are removed, and the trademarked words are extracted in the form of tokens. The trademark tokens extracted here are sets of English root words.

Step 2: Indexing Trademark Using the Hash Technique.

To reduce computational time during the search process, the features are indexed using a hashing technique. The hash indexing takes the trademark as the key index. It is then mapped to a list of trademark features from the database using a mapping function. The mapping function is designed so that the trademark similarity distance computation is performed only on the set of trademarks that consist of at least one of the terms in fs, i.e., the synonyms set belonging to the trademark query.

Step 3: Distance Computation between Trademarks

The distance computation is based on the similarity concept introduced in Tversky's contrast theory [11]. In this theory, Tversky defines the similarity between two objects as a function of unique and shared information about the object.

IV. ALGORITHM

In general, the retrieval algorithm consists of three main steps: 1) the feature extraction; 2) the hash indexing; and 3) the trademark similarity comparison measure. The feature extraction and the hash indexing are predominantly performed offline for indexing purposes, while the similarity computation is performed online. The proposed algorithm is competent of identify similar match up of trademarks from a database catalog and also, in a slightly different application scenario, such as an online application, finding trademarks similar to a particular trademark.
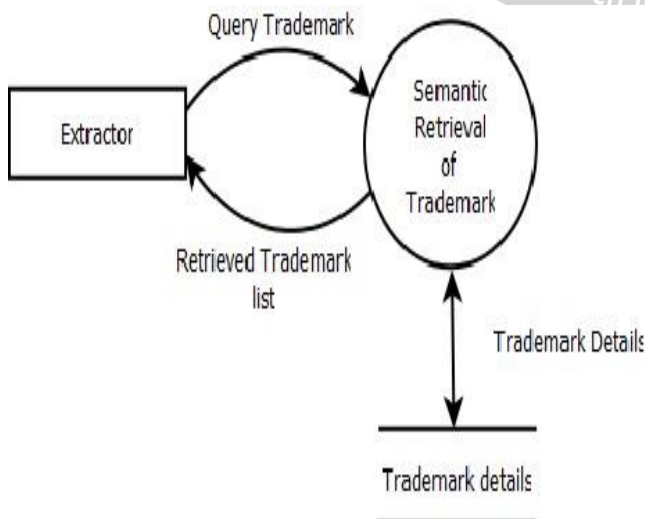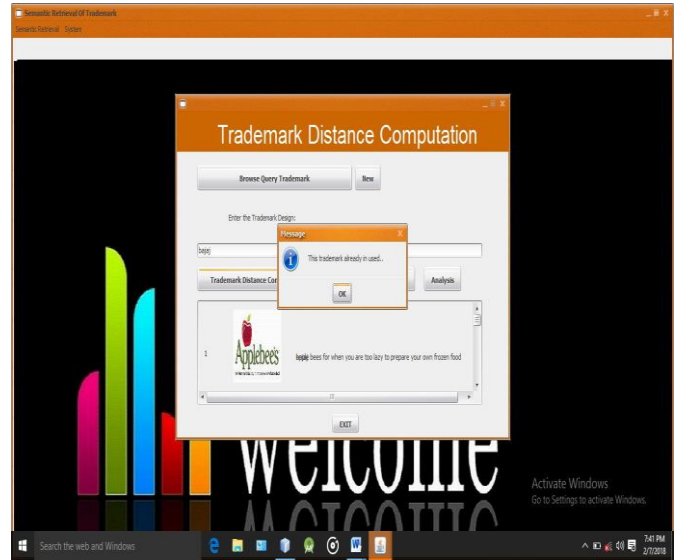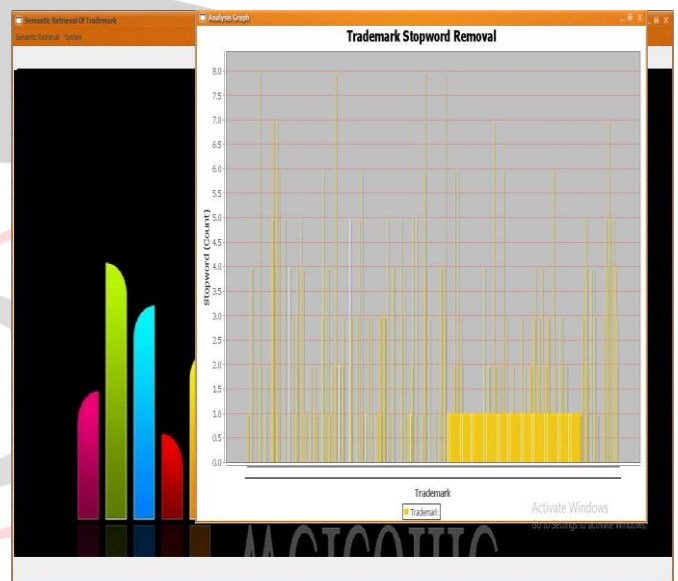
V. SYSTEM FLOW

Figure 2 Data Flow Diagram

Snap.1.Trademark Registered

Snap.2.Analysis Graph

VI. CONCLUSION

The work presented in this paper was inspired by the realization that regardless of the huge number of violation cases based on conceptual similarity, old information retrieval systems unable to handle this particular issue well. It is also provoked by the accepting that trademark match similarity, individual factors that add to the possibility of uncertainty, may be linked to the semantics of trademarks, i.e. their lexical meanings.

This paper contributes to the state-of-the-art by proposing a semantic algorithm to compare trademarks in terms of conceptual similarity. This algorithm fetches onward a completely innovative system that similarity assessment model in the domain of trademark retrieval. It utilizes NLP 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. The outcome of the research execute in this article validate that the assessment of trademarks based on their conceptual similarity can be conducted using linguistic resource.

Future work to improve the accuracy of the proposed semantic algorithm should include a study comparing the use of various lexical resources. as well, we are working on enhancing and extending this present approach to take account of retrieving trademarks with phonetic similarities and integrating their previous work on visual similarity with their new algorithms for conceptual and phonetic similarity.

ACKNOWLEDGMENT

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] Fatahiyah Mohd Anuar, Rossitza Setchi, and Yu-Kun Lai, "Semantic Retrieval of Trademarks Based on Conceptual Similarity," IEEE transaction on systems, man, and cybernetics: systems, vol. 46, no. 2, February 2016.

[4] Fatahiyah Mohd Anuar," A conceptual model of trademark retrieval based on conceptual similarity", 17th International Conference on Knowledge Based and Intelligent Information and Engineering Systems -KES2013

[5] H. Qi, K. Q. Li, Y. M. Shen, and W. Y.Qu, "An effective solution for trademark image retrieval by combining shape description and feature matching," Pattern Recognit., vol. 43, no. 6, pp. 2017-2027, 2010.

[6] M.-Y. Pai, M.-Y.Chen, H.-C.Chu, and Y.-M. Chen, "Development of a semantic-based content mapping mechanism for information retrieval," Expert Syst. Appl., vol. 40, no. 7, pp. 2447-2461, 2013.

[7] 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.

[8] S. A. Fadzli and R. Setchi, "Concept-based indexing of annotated images using semantic DNA," Eng. Appl. Artif. Intell., vol. 25, no. 8, pp. 1644-1655, 2012.

[10] F. M. Anuar, R. Setchi, and Y. K. Lai, "A conceptual model of trademark retrieval based on conceptual similarity," in Proc. 17th Int. Conf. Knowl. Based Intell. Inf. Eng. Syst., Kitakyushu, Japan, 2013, pp. 450–459.

[11] A. Tversky, "Features of similarity," Psychol. Rev., vol. 84, no. 4, pp. 327–352, 1977.