OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

SEMANTIC RETRIEVAL OF TRADEMARKS AND LOGOS BY DATA SIMILARITY

Ms. Sheetal Kusal[1], Mr. Prasad Gaikwad[2], Ms. Akanksha Bhat[3], Ms. Neha Malik[4]
Assistant Professor, Department of Computer Engineering, G.H.R.I.E.T Pune, India[1]
Students, Department of Computer Engineering, G.H.R.I.E.T Pune, India[2 3 4]

Abstract: **A trademark is a mark that you can use to recognize your business products or services from those of other vendors. It can be represented graphically in the form of any Symbol, logo, words etc. so, they need to be protected. The conceptual similarities among trademarks, which happens when more than two or more trademarks are similar. Trademarks are possessory words and images with high reputation they are main assets, often used as an application, which need infringement protection. The problems considered until infringement cases is the aspects, hypothetic and phonetic similarity of various trademarks. This paper focuses on an important aspect by proposing a conceptual similarity of trademarks and logos that can provide distance computation and suggestions of input retrieving conceptually similar trademarks and logos.**

The search and indexing technique developed uses similarity distance, which is derived using of similarity trademark. Propose a computational approach based on semantics that can be used to suggest the input of trademarks for conceptual similarity and to avoid the additional cost of protection to future infringement. A trademark retrieval system is performed with the massive number of semantic trademark of the conceptual similarity.

Keywords: Conceptual similarity, trademark infringement, trademark retrieval, trademark similarity, logo comparison.

.----------∴∴∴----------

I INTRODUCTION

The rapid development of simple ways has created new challenges in these regions for lots of companies who use the Internet to trade and employ trademarks as sell-out equipment. Trademarks, as prescribed by the European Office of Harmonization in the Internal Market (OHIM). They do insignificant intellectual property (IP) goods that permit good or service to be well validated to clients. Each year, many trademarks are registered and used. Trademarks are exclusive words or figures with advance reputational significance, used in commerce for comparison between products and services.

They allow products or tasks to be goods tenable and compared by traders. Searching for conceptually similar trademarks is a text retrieval problem. However, traditional text retrieval systems based on keywords and are not capable of retrieving conceptually related text. This limitation motivates research into semantic technology, which addresses this problem by using additional knowledge sources. Few common disservice outcomes from trademark infringement is lost income, low benefits, and need extra money of conservancy to stave off next infringement.

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 or logo. Infringement may occur when one party, the "infringer", uses a trademark which is identical or confusingly similar to a trademark owned by another party, in relation to products or services which are identical or similar to the products or services which the registration covers having an existing trademark look for systems as a general rule use text-based acts to get back technology. 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 difference of text files that part of the same domain, utilization same notations, or demonstration same consideration has been used broadly.

II LITERATURE SURVEY

The abstract comparison of text documents that share similar domain, categorical similar ideas, or use similar ideas has been studied. However, the abstract comparison of logos could be a distinctive drawback. as an example, logos are thought-about short texts [3]. They, therefore, need a new approach so as to spot the semantic similarities between logos. Most established methodologies for the semantic comparison of texts specialise in long texts [4]. However, thanks to the restricted variety of words in trademark texts, these methodologies don't seem to be applicable during this context, and thus, a new resolution is needed.

In addition, previous work addressing the difficulty of trademark similarity has targeted on visual comparison and analysis. The studies during this space are dominated by analysis on vision analysis and content-based data retrieval, similarly as developing systems capable of retrieving visually similar Trademarks [5]–[9].

In a recently launched search system, the OHIM provides associate possibility that enables users to go looking for Trademarks in numerous languages [10].

In U.K., the IP Office (IPO) provides search choices that are just like the OHIM search service, with a further possibility that searches for similar question strings [11]. Easy search tasks may fit well with ancient info systems. However, they are doing not work well once activity advanced tasks [12]. the previous causes ambiguity and results in the retrieval of spurious things, whereas the latter might cause a text containing relevant synonyms to not be retrieved, that conjointly results in poor performance.

The emergence of semantic retrieval technology was galvanized by the constraints of ancient keyword-based retrieval. Semantic retrieval employs external data sources, like ontologies, to beat the constraints of keyword-based systems [13]–[16]. Ontologies, that kind structural frameworks for organizing data, offer underlying domain-specific technical support, in conjunction with a theoretical basis for data illustration and organization [17].

Thus, these previous papers address the constraints of existing trademark retrieval systems, that presently use ancient text-based searches, by proposing a retrieval rule that retrieves emblems supported their abstract similarities.

There are four types of trademark images. These are the word-in mark, the device mark, the composite-mark and the complex mark (Kim & Kim, 1998). The word-in mark includes only characters or words while the device mark contains graphical or figurative shapes. The composite mark is a combination of these two types, and the complex mark is a composite mark which may have added color and visual effects. Psychological studies have shown that humans are more inclined to recognize and distinguish familiar objects based on their shape (Schietse, Eakins, & Veltkamp, 2007) [18].

For shape description, edge directions in WBS, contour curvature and centroid distance in TCS only represent the feature of each boundary point, without considering the relationship between adjacent boundary points. For feature matching, we have to get the weight value in WBS and the threshold value in TCS [19].

The local features do not change when each part of the image is rotated independently. As rotation invariant image features, moments such as Zernike moment and Legendre moment are known [6-8]. These are defined for the entire image as global features and are invariant to the rotation of the entire image [20].

III SYSTEM ARCHITECTURE OF PROPOSED SYSTEM

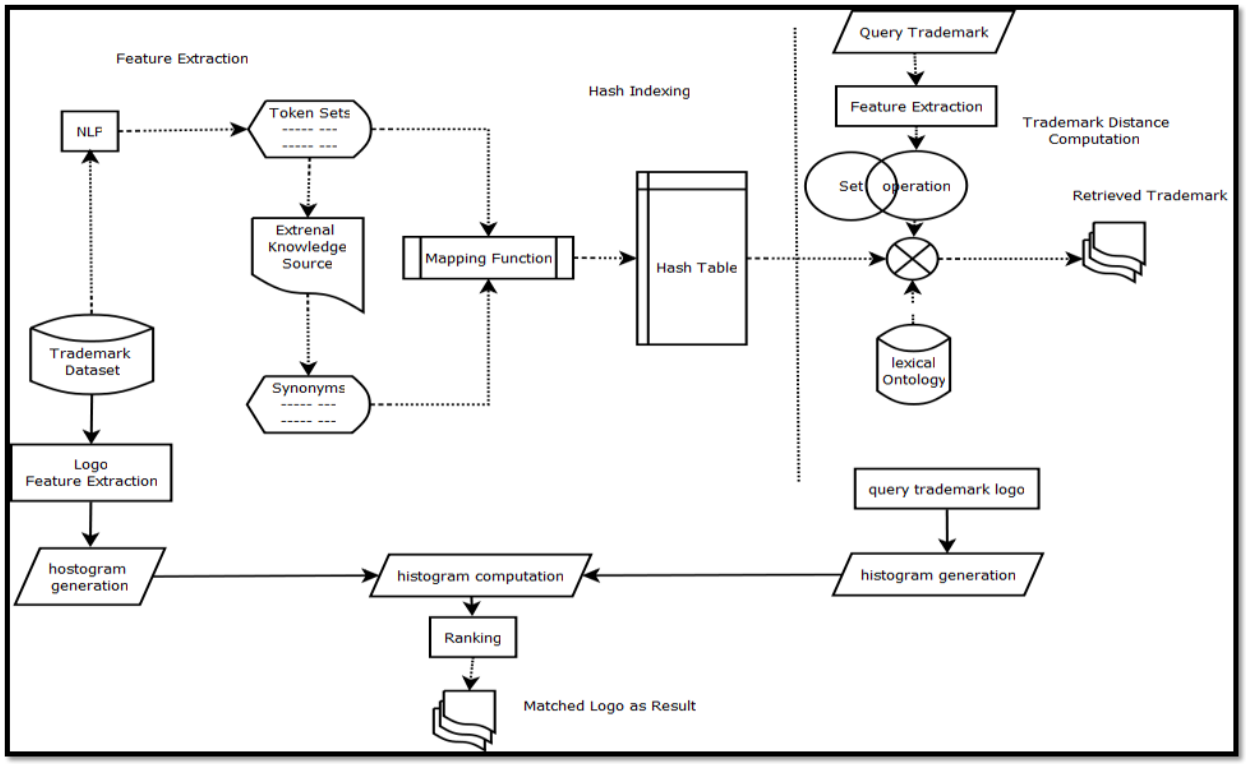
The proposed system is to make a retrieval of trademark hypothetical similarity to make them more accurate and more secure against the trademark infringement. Also, the systems are competent of retrieving the conceptual similarity of trademarks and manage the conventional data retrieval system. The proposed model can then be unified into a reflow system that considers the other two phases of similarity, sight and phonetic, and will then procedure a more extensive trademark comparison. The system used to proportion trademarks for conceptual similarity. Finding for conceptually same trademarks is a text retrieval problem. The system defines the nearly string matching which is used to text searches.

The proposed algorithm uses techniques for the word similarity gap of that method, which was already use from the Word Net ontology and it’s together with a new trademark comparison measure. In data retrieval, data is per performed for exhibit classifications, as well as realizing potentially useful information from documents.

The analysis of the trademarks is needed to comprehend the main of conceptual similarities coming from different factors. The focuses on this main fact by proposing a hypothetical model of the comparison process, purposed at retrieving conceptually similar trademarks. The hash indexing accept the token key and synonym key to pre-processing and use the indexing in that key and create a new trademark for the user, its similar to that user requirement trademark. The feature extractions are defined the token and synonyms. The excerption compatible conceptual features, which are then used to proper manage the database. The spelling corrector corrects any spelling mistakes in the trademark text, and can be adapted from

any existing spell checker. The spelling corrector is mainly work for conceptual similarity of trademark that is not generating the wrong trademark in the system.

The stop words remover removes frequent words e.g. no, and, the, etc. for improve the efficiency of the trademark. The pre-processing is made a require changes then that trademark text provide the token for unique identifications and also provided synonyms word from lexical resource for create new trademark as user conceptual similarity. In point of extractions, the provider of these steps two features: 1) Token set and 2) Synonym set. The token set is defined as a particular word has a token set for used in the trademark retrieval system. After stop word removal, every single word has a token set.



The synonym set is defined as direct token set is extracted of every word has synonyms used word net. This is use to provide a every word of token has a number of synonyms find out using word net, if not then keep as it token word.

To remove hypothetic time throughout the find process, the factors are listed using a hashing technique. The hash indexing accepts the trademark as the key index. The feature extraction process stores an additional set of array features in the first feature vector, i.e. a set of synonyms and antonyms corresponding to the query tokens. The hash table is use the hash indexing for provide accurate timing for searching the process for conceptual similarity of trademark. In hash table use the token key as well as synonyms key for unique identification. The token key and synonym key as used indexing for provide new retrieval trademark. The retrieval trademark list is content of number of trademarks of the conceptual similarity. Using retrieval algorithm for the similarity between two words. The hash indexing develops the new trademark using token key and synonym key in the hash table.

A trademark reflow technique using the proposed retrieval algorithm is evolved, and the algorithm is tested on conceptual similarity. The retrieval trademark list is stored in database for next future trademark use in the next trademark retrieval concepts. To remove extra required time throughout the find procedure, the factors are listed using a hashing technique. The hash indexing is taken the trademark as the key index. Through trademark retrieval process user can enter a text which he wants to trademark. If trademark is already exist in system then it sent to trademark matching and return the similar documents to the user. If trademark is not existed in system then trademark is stored in database. The return document is send to user is the use lexical resource and apply the hash indexing to that trademark for create new trademark to get the user.

Histogram: 1. Learning Stage: The histogram represents which kind of features is dominant and which is less dominant. We regard the histogram as the image descriptor. The histograms are calculated for learning images and

2. Retrieval Stage: In the retrieval stage, first, rotation invariant features are extracted from the query image. Then, the quantization and the histogram computation are performed in the same way as in the learning stage. Next, the similarity between the histogram obtained from the query image and each histogram in the dictionary is calculated. The similarity is defined as the absolute difference between two normalized histograms of the query image and the ith histogram in the dictionary. Finally, the targeted images are ranked according to the similarity.

IV CONCLUSION

The work was motivated by increasing of fraud cases best as, data similarities, where information retrieval system does 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 hypothetical model of retrieval trademark depends on hypothetical similarity. The main model language processing technology, data paths and lexical resources to calculate hypothetical 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 problem. The system reforms on all ready trademarks find a system by legislation a implementing of rectification the find to hypothetical same trademarks. The system employs natural language processing techniques, knowledge sources and a lexical resource compute conceptual similarity between trademarks. Also confirm that the comparison of trademarks in terms of conceptual similarity. In future work to improve the precision of the proposed semantic algorithm should include a study comparing the use of various lexical resources.

REFERENCES

[1]F. M. Anuar, Yu-Kun Lai, R. Setchi,"Semantic Retrieval of Trademarks Basedon Conceptual Similarity." IEEE Transactions on Systems, Man, and Cybernetics: Systems IEEE permission, 2015.

[2] F. M. Anuar, R. Setchi, and Y. K. Lai, "Trademark image retrieval using an integrated shape descriptor,"Expert Syst. Appl., vol. 40, no. 1, pp. 105-121, 2013.

[3] F. M. Anuar, R. Setchi, and Y. K. Lai, "Trademark image retrieval using an integrated shape descriptor," Expert Syst. Appl., vol. 40, no. 1, pp. 105–121, 2013.

[4] 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.

[5] C. H. Wei, Y. Li, W. Y. Chau, and C. T. Li, "Trademark image retrieval using synthetic features for describing global shape and interior structure," Pattern Recognit., vol. 42, no. 3, pp. 386–394, 2009.

[6] 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.

[7] J. P. Eakins, J. M. Boardman, and K. Shields, "Retrieval of trademark images by shape feature—The ARTISAN project," in Proc. IEEE Colloq. Intell. Image Databases, London, U.K., 1996, pp. 9/1–9/6.

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

[11] Trade Mark Search. [Online]. Available: http://www.ipo.gov.uk, accessed Dec. 10, 2013.

[12] R. Bhagdev, S. Chapman, F. Ciravegna, V. Lanfranchi, and D. Petrelli, "Hybrid search: Effectively combining keywords and semantic searches," in The Semantic Web: Research and Applications (Lecture Notes in Computer Science), vol. 5021. Berlin, Germany: Springer, 2008, pp. 554–568.

[13] 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.

[14] 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.

[15] 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.

[16] R. Setchi, Q. Tang, and I. Stankov, "Semantic-based information retrieval in support of concept design," Adv. Eng. Inf., vol. 25, no. 2, pp. 131–146, 2011.

[17] V. C. Storey, D. Dey, H. Ullrich, and S. Sundaresan, "An ontology-based expert system for database design," Data Knowl. Eng., vol. 28, no. 1, pp. 31–46, 1998

[18] Fatahiyah Mohd Anuara, Rossitza Setchia, Yu-kun Laib, "Trademark image retrieval using an integrated shape descriptor", 2012
[19] Heng Qia, Keqiu Lia, Yanming Shena, Wenyu Qub, "An effective solution for trademark image retrieval by combining shape description and feature matching", 2010
[20] Masafumi Miyazaki, Kyoko Miyazaki, Keisuke Tod, "A Similar Trademark Retrieval System Based on Rotation Invariant Local Features", 2016