OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

SURVEY ON SIMILARITY BASED RETRIEVAL OF TRADEMARKS AND LOGOS BY DATA SIMILARITY

Neha Shingare[1], Priyanka Jadhav[2], Yogita Bhagat[3], Pranita Shinde[4], Prof. Anuja Bharate[5]

JSPM's, Imperial College Of Engineering and Research, Wagholi Maharashtra, India[1,2,3,4,5]

nehashingare08@gmail.com[1], priyajadhav2930@gmail.com[2], yogitabhagat95@gmail.com[3], shindepranita75@gmail.com[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 protection. The conceptual similarities among trademarks, which happens when more than two or more trademark similar. Trademarks are possessory words and images with high reputation they are main assets, often used as a 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 important aspect by proposing a conceptual similarity of trademarks that can be provide distance computation and suggestions of input retrieving conceptually similar trademarks. 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 performing with the massive number of semantic trademark of the conceptual similarity.

Keywords: *Conceptual similarity, trademark infringement, trademark retrieval, trademark similarity.*

---∴∴∴---

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 well or service to be well validated to clients. Each year many trademarks registered and used that outlet. Trademarks are exclusive words or figures with advance reputational significance, used in commerce to 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 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 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 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 existence 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 different of text files that part of same domain, utilization same notations, or demonstration same consideration has been used broadly.

II LITERATURE SURVEY

In this paper [3], the recent trademark reflow system of working with reformed reflow execution for the unification of global and local expositors. The global expositors are using the Zernike moment's coefficients and the local expositors are the edge-gradient co-occurrence matrix, defines as outline data that means it's mainly significance in human cognition of estimation equality. The defined reflow system is tested use the standard MPEG-7 shapes. The results reformation in the case of the MPEG-7 shape databases. The bonding during two proximate factors is hold on by usage the co-occurrence matrix on incline data. The research in the round of offered a novel system for trademark reflow that increase the execution.

Author proposed [9], 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 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.

Author introduces [5] 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's 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 indurations probe suggests that the method is efficient to many forms and define nations with accurate purity.

This paper presents [6], 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. A semantic-based content mapping mechanism uses files different keywords as the input, which substances the semantic characteristics and fabrication of the documents.

This paper [7] the problems define during infringement litigation is the visible, hypothetic and phonetic similarity of different trademarks. This is focuses on this important fact by defining a hypothetic model of the comparison process, target at retrieving hypothetic similar trademarks. The proposed model normal language accessing and semantic technology to get the hypothetic similarity between trademarks. Proposes a hypothetic model of trademark retrieval based on hypothetic similarity. The proposed model improves on already trademark finding models by providing find to hypothetically related trademarks.

III PROBLEM DEFINATION

As deep Trademarks violation is an aspect of IP delinquency that hegemony to serious financial issue that is it trademark infringement is lost revenue. The search and indexing technique developed uses similarity distance. Hence, the concept of similarity has become well understood in trademark infringement litigation. A successful one of the aspects of similarity assessed during trademark analysis, which is conceptual and logo similarity.

IV WORK OF TRADEMARK SYSTEM

The proposed retrieval algorithm and histogram algorithm are based on a conceptual and histogram model respectively. The trademark comparison process developed in It provides a bird's eye view of trademark comparison based on conceptual and logo similarities. This system extends the conceptual model by developing and evaluating a semantic and histogram algorithm for trademark retrieval based on conceptual similarity and histogram. The proposed algorithm employs NLP techniques and the word similarity distance method, which was derived from the Word Net ontology, together with a new trademark comparison measure. Word Net is employed in this algorithm due to its lexical relationships, which mirror human semantic organization, and because it has also been proven successful in many previously developed works. The trademark comparison measure is derived from the Tversky contrast model, a well-known model in theory of similarity.

The propose 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 and logo similarity of trademarks and manage the conventional data retrieval system. The proposed model can then be unified into a re-flow 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. System also use histogram for the problem of partial similarity, we determine to use a similarity measure based on the histogram. Because histogram of any partial rectangular region can be calculated effectively.

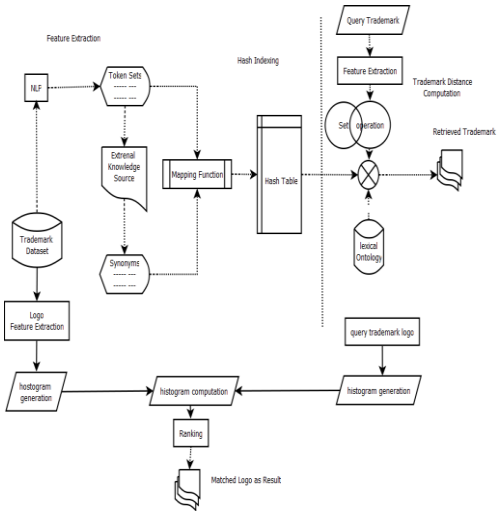

Figure 1: System Architecture

The analysis of the trademarks is needed to comprehend the main of conceptual and logo 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. 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.

V MATHAMATICAL MODEL

A. Set Theory

Let, S be a system, S= (F, ft, fs, H, Mf, Ht, Q, D, Tr) where,

1. Feature Extraction
   F= {ft, fs},
   F is set of feature extraction.
   ft is set of token set.
   fs is set of synonyms set.
2. Hash Indexing
   H= {Mf, Ht},
   H is set of Hash Indexing.
   Mf is set of mapping function.
   Ht is set of hash table.
3. Query Trademark
   Q= {Q1, Q2,...,Qn} ,
   Q is set of query trademark.
   - Distance Computation and suggestions
     D= {D1, D2,..., Dn},
     D is set of distance computation and suggestions.
   - Retrieval list
     Tr= {Tr1, Tr2,...,Trn},
     Tr is set of retrieval list.

B. Mathematical model for proposed system

For trademark distance computation,

$$sim(Q,T) = \left|\frac{Qf_t \cap Tf_t}{Qf_t \cup Tf_t}\right| + \frac{|Qf_s \cap Tf_t|}{D} + \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\max\left(wordsim(x_{i,y_i})\right)}{|Qf_t \backslash Tf_t|.|Tf_t \backslash Qf_t|}$$

$Qf_t \& Qf_s$: Token set and synonyms set of the query.

Tf_t: Token set of one of trademark from database.

C. Retrieval Algorithm

Algorithm 1 Pseudo code of proposes system is:

1: define ft as the token set of a trademark
2: define fs as a set of synonyms that correspond to the token set
3: define as a list of unique token extracted from the database
4: for each trademark in the database
5: do (extract ft, extract fs)
6: for each token in ft
7: if (token does not exist in ft)
8: update token into ft
9: define hash table as index table that maps token to trademarks in the database
10: define Q as the query of trademark

11: if (Query is present in database)
12: then (display query suggestions of trademark)
13: else (Use this as a trademark)

VI CONCLUSION

The work is 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 sys-tem work, conceptual and logo 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. We are also using a new similar trademark retrieval process. As for the problem of partial similarity, we determine to use a similarity measure based on the histogram. Because histogram of any partial rectangular region can be calculated effectively.

ACKNOWLEDGEMENT

It is pleasure for me to thank lot of people who have support me in completing to this paper in different ways. Starting with, I would like to thank my Guide, **Prof. Anuja Bharate** for her support during the entire work. Her support for helping me to get helpful content in the topic. I am thankful to HOD, Principal, teachers and collage to providing us the facilities in the department and their guidance.

REFERENCES

[1] Dhruvil Shah, Prathmesh Kudale, Prasad Shirwadkar, Samuel Jacob, Iot Based Air and Sound Pollution Supervising System, IOSR Journal of Engineering, 2018.
[2] Arushi Singh, Divya Pathak, Prachi Pandit, Shruti Patil, Prof. Priti . C. Golar, IOT based Air and Sound Pollution Monitoring System, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2017.
[3] Sindhu.K.G, Shruthi.H, Sumanth.M.B, Vijayashree.H.M, Ayesha.A.P, IOT Based Air and Noise Pollution Monitoring System, International Journal of Innovative Research in Science, Engineering and Technology, 2018.
[4] Ms. Sarika Deshmukh, Mr . Saurabh Surendran, Prof. M.P. Sardey, Air and Sound Pollution Monitoring System using IoT, International Journal on Recent and Innovation Trends in Computing and Communication, 2017.
[5] P. Sai Chandana, K. Sreelekha, A. Muni Likith Reddy, M. Anil Kumar Reddy, R. Senthamilselvan, IOT Air And Sound Pollution Monitoring System, International Journal on Applications in Engineering and Technology, 2017.
[6] Anushka Sharma, Vaishnavi Varshney, Roopank Maheshwari, Upasana Pandey, IOT Based Air And Sound Pollution Monitoring System, International Research Journal of Engineering and Technology, 2018.
[7] Lalit Mohan Joshi, Research paper on IOT based Air and Sound Pollution Monitoring System, International Journal of Computer Applications, 2017.