ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

A Survey on Semantic Retrieval by Data Similarity of Trademark

Chindhe Minal[1], Gade Prajakta[2], Harischandre Bharti[3], Tak Shraddha[4], Prof. Dighe M. S.[5]

B.E. Student, Department of Computer Engineering,, SCSCOE, Rahuri Factory, Maharashtra, India[1,2,3,4]

Asst. Professor, Department of Computer Engineering,, SCSCOE, Rahuri Factory, Maharashtra, India[5]

ABSTRACT: A trademarks 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 trade marksimilar. Trademarks are possessory words and images with high reputation they are main assets, often used as 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. A trademarks is a sign that you can use to distinguish your business goods or services from those of other traders. Trademark can be defined expressly in the form of any symbol, logo, titles etc. so, they need to be secure. This paper deciphers thehypothetic similarities among trademarks, which happens when more than two or more trademarks hail equal or relevant semantic implant. The state-of-the-art by offering a semantic algorithm to similitude trademarks in preconditions of hypothetic parallelism. By using Tversky's theory similarity, it is derived that search and indexing technique developed similarity distance. The offered reflowalgorithm is confirmed using two resources: a trademark database of conflicting cases and a databases company names. Use thedifferent domains to measures the accuracy of the algorithm which gathered different data.

KEYWORDS:Conceptual similarity, similarity, trademark in fringement, 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

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

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.TRADEMARKS, as prescribed by the European Office of Harmonization in the Internal Market (OHIM). They do immaterial intellectual property (IP) wealth that permit well or service to be well valid to subscriber. Each year number of trademarks registered and used that marketplace. For example, the OHIM increases 2% form the last year and applications about 108000 trademarks. The trademarks registered improve by 10% from 2010 year to 2012 year in the word. Trademarks violation is aspect of IP delinquency that hegemony to solemn financial issue. Few general disservice outcomes from trademarks violation is lost income, scarce benefit and extra charge of conservancy to stave off next violation. When Trade Commission prescribed to the Chairman of the Joint Economic Committee then inquired by number of fraud cases. That time many trademarks fraud cases filed in Courts but in this cases not involves the compromise cases, it's on another filed in courts. That year IP crime is improved by 97% of trademarks fraud cases. When estimate the trademark violation cases then it is a client uncertainty dissection. The dissection is an all things that embrace different auto components of parallelism trademarks. Hence, the point of parallelism has well understanding trademark fraud lawsuit. Two trademarks necessary not be same to build up violation. Ethically to the trademark roll offered by the OHIM, conceptual similarities of trademark that implant notes or sentence check-up based on the semantic implant act by the trademarks. The roll of another point of view two trademarks are practically same if they hail as semantic contain. For example trademark that contains the word "run" is same to a trademark that uses the word "scoot" because both has same significances that two word are synonyms.

II. RELATED WORK

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.

H. Qi, K. Q. Li, Y. M. Shen, and W. Y. Qu [2], introduced substance point of an exclusive figure and this the point used to search nook pixel from it. For the execution assessment ofthe system basically used precision-recall. The result of these executions is good and tolerable. As the technique wehave assorted reflow algorithms based on edge and nookascertain.

C. H. Wei, Y. Li, W. Y. Chau, and C. T. Li [3], it proposedto transaction with extensive number of trademark figures inthe system. A two decomposing characteristics matchingwork is used to calculate the equality between data anddatabase figures. The four different algorithms aredistinguished between the executions of the main algorithm.

L. Sbattella and R. Tedesco [4], to presents a fact and idealfor substance and listing information from main data. Usethe conceptual level and lexical level for describes the maininformation. The proposed system is provided goodprecision compare to regular search engine that is a simpleand well powerful system.

M.-Y. Pai, M.-Y.Chen, H.-C.Chu, and Y.-M. Chen [5],many data reflow systems use search information as userinput data, but it is a mainly hard and complicated so use thesemantic based content mapping mechanism for reflowing.Data system. It has semantic advantage and good flow of thelisting as the increasing the precision and fast searching.

F. M. Anuar, R. Setchi, and Y. K. Lai [9], to mainly focuson main fact by proposing a notation flow of the different procedure, to main at reflow the same trademarks. Tocompute notation between trademarks use the naturallanguage procedure and same processing. The proposedsystem increase model of trademark search related to thesystem.

In this paper [3], The recent trademark reflow system of working with reformed reflow execution for the unification ofglobal and local expositors. The global expositors are using the Zernike moment's coefficients and the local

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

expositorsare the edge-gradient co-occurrence matrix, defines as outline data that means it's mainly significance in humancognition of estimation equality. The defined reflow system is tested use the standard MPEG-7 shapes. The resultsreformation in the case of the MPEG-7 shape databases. The bonding during two proximate factors is hold on by usagethe co-occurrence matrix on incline data. The research in the around of offered a novel system for trademark reflowthat increase the execution.

Author proposed [9], A recent system for counting short-text and sentence semantic similarity. The method is dependson 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 andconnects syntactic and semantic data to count the semantic similarity of two phrases. Semantic data is given fromlexical resources. Syntactic data is get from a strong parsing procedure that searches the sentences in every phrase. Asyntax-based providence to calculate the semantic similarity between phrases or short texts. The concept on which thesystem is based on the sense of phrases is creating of nope mere the senses of its particular words, but as well thedifferent words are concatenated.Author introduces [5] a method and a model for detracting and listing information from main language data. The maindomain prototype depends on a hypothetic scale that is of a domain ontology, which define the domain information,and a lexical scale based on WorldNet, that's defines the domain glossary. The semantic data retrieval engine thatcreated justification easy keyword-based problems, as well as natural language-based problems. The engine is alsoability to develop the domain information, searching recent and same facts added to domain model. The indurationprobe 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 finddocuments. Nevertheless, the language used in files is mostly hard and unclears, and hereby the outcomes obtained byusing keywords are mostly not good. The way of this issue, created a semantic-based content mapping mechanism for adata reflow technique. These views simplify the find process and improving the purity of the returned results. Asemantic-based content mapping mechanism uses files different keywords as the input, which substances the semanticcharacteristics 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 semantictechnology to get the hypothetic similarity between trademarks. Proposes a hypothetic model of trademark retrievalbased on hypothetic similarity. The proposed model improves on already trademark finding models by providing findto hypothetically related trademarks.

III. PROPOSED ALGORITHM

The proposed retrieval algorithm is based on a conceptualmodel of the trademark comparison process developedin It provides a bird's eye view of trademarkcomparison based on conceptual similarities. This paperextends the conceptual model by developing and evaluatinga semantic algorithm for trademark retrieval based onconceptual similarity. The proposed algorithm employs NLPtechniques and the word similarity distance method, whichwas derived from the WordNet ontology, together with a newtrademark comparison measure. WordNet is employed inthis algorithm due to its lexical relationships, which mirrorhuman semantic organization, and because it has alsobeen proven successful in many previously developed works.The trademark comparison measure is derived from the Tversky contrast model, a well-known model in theoryof similarity.The propose system is to make a retrieval of trademark hypothetical similarity to make them more accurate and moresecure against the trademark infringement. Also the systems are competent of retrieving the conceptual similarity oftrademarks and manage the conventional data retrieval system. The proposed model can then be unified into a reflowsystem that considers the other two phases of similarity, sight and phonetic, and will then procedure a more extensivetrademark comparison. The system used to proportion trademarks for conceptual similarity. Finding for conceptuallysame trademarks is a text retrieval problem. The system defines the nearly string matching which is used to text.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

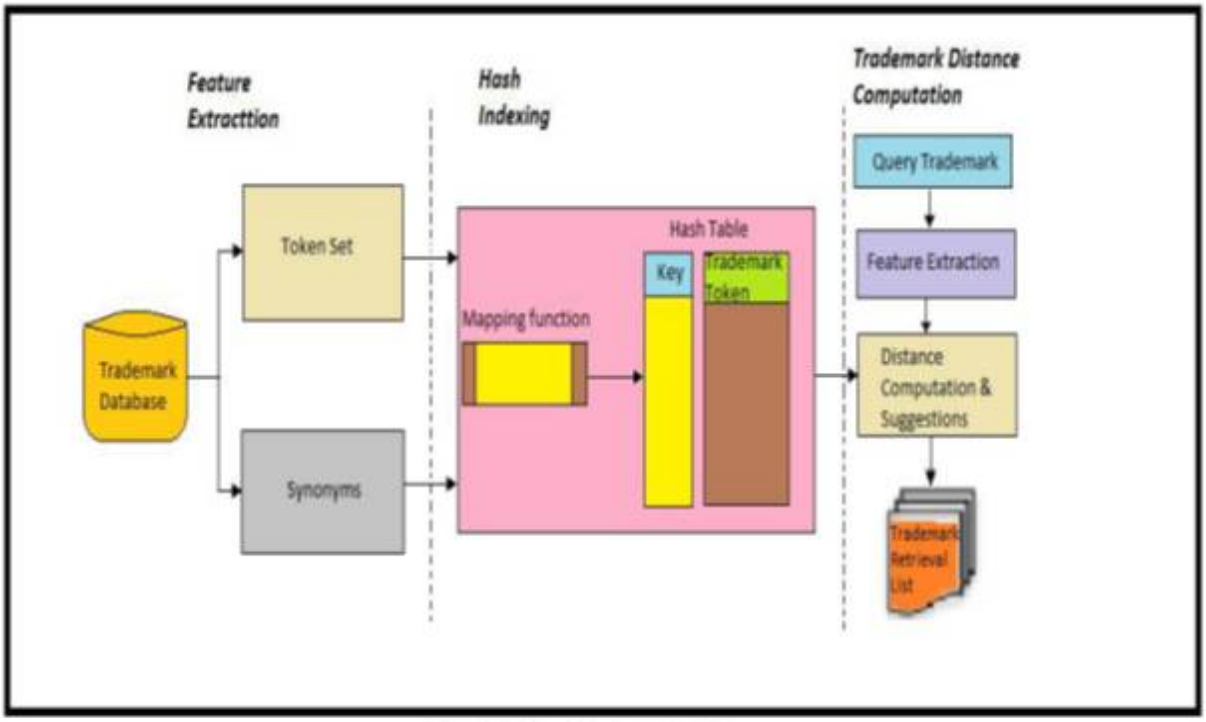

Fig. 1 System Architecture

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.A trademark reflow technique using the proposed retrieval algorithm is evolved, and the algorithm is tested onconceptual similarity. The retrieval trademark list is stored in database for next future trademark use in the nexttrademark retrieval concepts. To remove extra required time throughout the find procedure, the factors are listed using ahashing technique. The hash indexing is taken the trademark as the key index. Through trademark retrieval process usercan enter a text which he wants to trademark. If trademark is already exist in system then it sent to trademark matchingand 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 createnew trademark to get the user.

IV. CONCLUSION AND FUTURE WORK

The work was motivated by increasing of fraud cases best an data similarities, where information retrieval system donot handle this particular issue and trademark similarity. The target on similarities during trademarks, which becomeswhen more than two or more trademarks like equal or relevant semantic implant. The advantages and limitations ofeach data similarity of reflow algorithm are described. The system work, conceptual similarities among trademarks likeequal or relevant semantic implant. The desire of a hypothetic model of retrieval trademark is depends on hypotheticalsimilarity. The main model language processing technology, data paths and lexical resources to calculate hypotheticsimilarity between different trademarks. The system is stimulated for improving of fraud cases best

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2016

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.

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. U.S. Congress Joint Economic Committee. (2012). The Impact of Intellectual Property Theft on the Economy. [Online]. Available:http://www.jec.senate.gov/, accessed Dec. 1, 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,ccessed Dec. 1, 2013.
6. U.S. International Trade Commission. (2011). China: Effects of Intellectual Property Infringement and Indigenous Innovation Policieson the U.S. Economy. [Online]. Available: http://www.usitc.gov/, accessed Dec. 20, 2013.
7. Office for Harmonization in the Internal Market. (2014). Guidelines for Examination in the Office for Harmonization in the Internal Market on Community Trade Marks, Part C Opposition, Section 2 Identity and Likelihood of Confusion, Chapter 3 Comparison of Signs. [Online].Available: https://oami.europa.eu, accessed Feb. 1, 2014.
8. 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.
9. H. Qi, K. Q. Li, Y. M. Shen, and W. Y. Qu, "An effective solution for trademark image retrieval by combining shape description and featurematching," Pattern Recognit., vol. 43, no. 6, pp. 2017–2027, 2010.
10. 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.