Semantic Retrieval of Trademarks Based on Conceptual Similarity

Fatahiyah Mohd Anuar, *Student Member, IEEE*; Rossitza Setchi, *Senior Member, IEEE;* and Yu-Kun Lai

***Abstract*— Trademarks are signs of high reputational value. Thus, they require protection. This paper studies conceptual similarities between trademarks, which occurs when two or more trademarks evoke identical or analogous semantic content. The paper advances the state-of-the-art by proposing a computational approach based on semantics that can be used to compare trademarks for conceptual similarity. A trademark retrieval algorithm is developed that employs natural language processing techniques and an external knowledge source in the form of a lexical ontology. The search and indexing technique developed uses similarity distance, which is derived using Tversky's theory of similarity. The proposed retrieval algorithm is validated using two resources: a trademark database of 1,400 disputed cases and a database of 378,943 company names. The accuracy of the algorithm is estimated using measures from two different domains: the R-precision score, which is commonly used in information retrieval, and human judgment/collective human opinion, which is used in human-machine systems.**

***Index Terms*— Conceptual similarity, similarity, trademark infringement, trademark retrieval, trademark similarity.**

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 United States, 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 US climbed by 10% from the 2010 fiscal year to the 2012 fiscal year [2].

Trademark infringement is a form of intellectual property crime that may lead to serious economic problems. In general, IP-intensive companies make twice as many sales as non-IP-intensive companies. In the United States, these companies contribute over one-third of the annual gross domestic product [3]. Some major damage resulting from trademark infringement is lost revenue, lower profits, and the additional cost of protection to avoid future infringement. In a statistic provided by the United States 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 3,400 trademark infringement cases were filed in US 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]. In another investigation, conducted in 2011 by the US International Trade Commission, it was found that trademark infringement is the most common form of IP crime in the fastest growing economy in the world: China [6]. The same investigation also revealed that US-based companies lost between $1.4 billion and $12.5 billion in 2009. In fact, between 2002 and 2011, the average annual increase in trademark litigation cases was 39.8%.

A compulsory analysis required by both European law and US legal practice [4, 7] 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.

Hence, the concept of similarity has become well-understood in trademark infringement litigation. 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. Two trademarks need not be identical to constitute an infringement. Moreover, similarity, in the context of trademarks, is also not binary but a matter of degree. The rule of thumb is that the higher the degree of similarity between the trademarks, the more likely it is that they will cause confusion. This paper addresses one of the

This Manuscript received on 31st May 2014; revised on 11th November 2014; accepted on 14th March 2015.

F. Mohd Anuar is a PhD student at Cardiff University, School of Engineering, CF24 3AA, Cardiff, UK. (e-mail: MohdAnuarF@cardiff.ac.uk).

R. Setchi is a Professor in Cardiff University, School of Engineering, CF24 3AA, Cardiff, UK. (e-mail: Setchi@cardiff.ac.uk).

Y-K. Lai is a lecturer in Cardiff University, School of Computer Science and Informatics, CF24 3AA, Cardiff, UK. (e-mail: Yukun.Lai@cs.cardiff.ac.uk).

aspects of similarity assessed during trademark analysis, which is conceptual similarity.

According to the trademark manual [7] produced by the Office of Harmonization in the Internal Market, a European Union agency responsible for registering trademarks and designs for all European countries, the conceptual similarity of trademarks that contain words or phrases is examined based on the semantic content portrayed by the trademarks. The manual further explains that two trademarks are conceptually similar or identical if they evoke identical or analogous semantic content. For example, a trademark that contains the word *'quick'* is similar to a trademark that uses the word *'fast'* because both evoke similar meanings (i.e., the two words are synonyms). Conceptual similarity also exists between the words *'hour'* and *'time'*. Although the two words are not synonyms, they are semantically related. Such a similarity comparison requires external knowledge sources in the form of dictionaries or encyclopedias, as suggested in the manual.

The conceptual comparison of text documents that share similar domain, use similar concepts, or express similar ideas has been studied extensively. However, the conceptual comparison of trademarks is a unique problem. For instance, trademarks are considered short texts [8]. They therefore require a new approach in order to identify the semantic similarities between trademarks. Most established methodologies for the semantic comparison of texts focus on long texts [9]. However, due to the limited number of words in trademark texts, these methodologies are not applicable in this context, and thus, a new solution is required.

In addition, previous work addressing the issue of trademark similarity has focused on visual comparison and analysis. The studies in this area have been dominated by research on vision analysis and content-based information retrieval (CBR), as well as developing systems capable of retrieving visually similar trademarks [10-14]. Although the amount of work and the outcomes have been encouraging, these approaches are mainly limited to trademarks with figurative marks and only cover one-third of the similarity criteria required in the assessment, i.e., the visual aspect. Additionally, as shown by the statistics of registered trademarks in five European countries, only 30% of all trademarks employ logos as their proprietary marks [15]; this leaves the remaining 70% still insufficiently researched.

The conceptual comparison of trademark words and phrases is therefore a new problem in the domain of trademark retrieval. It requires a cross-disciplinary approach involving natural language processing (NLP) and external knowledge sources (i.e., dictionaries or thesauri), which to the best of the authors' knowledge, have not been adequately studied until now. Hence, this paper provides a mechanism via which to compare the conceptual aspects of trademarks by proposing a trademark retrieval algorithm based on their conceptual similarity. The proposed algorithm employs a knowledge source in the form of a lexical ontology that is used together with Tversky's set similarity theory to retrieve conceptually similar trademarks. The proposed algorithm is then tested on two databases, a database of 1,400 disputed trademark cases from 1998–2012 and a company name database comprised of 378,943 names.

The rest of the paper is organized as follows. The next section provides an overview of related work. It discusses existing trademark search systems, the limitations of traditional information retrieval, the strengths of semantic retrieval, the lexical ontology employed, and existing word similarity measures. The proposed trademark retrieval algorithm is then discussed in Section III. Section IV describes the experimental setup. The results of the experiment, together with discussions, are provided in Section V, and Section VI concludes this study.

II. Related Work

A. Existing Trademark Search Systems

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 [16]. This newly upgraded system also provides advanced search options that offer three search types: word prefix, full phrase, and exact match. The word prefix mode returns trademarks with a prefix that matches the query. The full phrase mode finds trademarks with terms that include the query input, and the exact match returns trademarks that match the query input exactly.

In the United Kingdom, the Intellectual Property Office (IPO) provides search options that are similar to the OHIM search service, with an additional option that searches for similar query strings [17]. The system employs an approximate string-matching technique, along with several pre-defined criteria, such as the number of similar and dissimilar characters in the words and the word lengths, to retrieve similar trademarks. Approximate string matching is a commonly used algorithm that computes the similarity between two strings using edit distance, which is derived based on the number of insertion, deletion, and substitution operations that would be required to make the two strings identical. For example, the word string pair 'come' and 'some' requires only one substitution operation. The fewer operations required to make the strings identical, the more similar they are.

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 pre-defined 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. However, they do not work well when performing complex tasks [18]. For example, in the case of text retrieval, the effectiveness of keyword-based search suffers from two main issues related to polysemy (i.e.,

words with multiple meanings) and synonymy (several words with the same meaning). The former causes ambiguity and leads to the retrieval of spurious items, while the latter may cause a text containing relevant synonyms not to be retrieved, which also leads to poor performance.

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 [19-22]. Ontologies, which form structural frameworks for organizing information, provide underlying domain-specific technical support, together with a theoretical basis for knowledge representation and organization [23]. For example, a lexical ontology contains lexical knowledge source relationships between its entries, as defined by lexicons. In text retrieval, this allows for the semantic processing of document content, which cannot be achieved through traditional text mining.

Thus, this paper addresses the limitations of existing trademark retrieval systems, which currently employ traditional text-based searches, by proposing a retrieval algorithm that retrieves trademarks based on their conceptual similarities.

B. Lexical Knowledge Sources and Semantic Similarity

Retrieving conceptually similar trademarks requires semantic interpretation, which can be realized using lexical knowledge sources. Lexical knowledge sources include lexicons, thesauri, and dictionaries that have been semantically formalized in accordance with the lexical meanings of the words. The lexical knowledge source employed in this study is WordNet, a large electronic lexical database of English language words. This freely available database is one of the most frequently cited lexical resources in NLP literature, with many applications in a wide range of tasks.

Developed by the Cognitive Science Laboratory at Princeton University, USA, WordNet was constructed based on psycholinguistic theories that model human semantic organization. Nouns, verbs, adjectives, and adverbs are grouped into sets of cognitive synonyms that act as building blocks known as synsets [24]. Each synset represents a distinct concept and is linked by lexical relationships, such as synonymy, antonymy, hyponymy, and meronymy [25]. Additionally, each synset also contains a short definition, or gloss, which in most cases includes at least one sentence illustrating the usage of the synset members. To date, WordNet has been successfully established in over 30 languages (e.g., Dutch, Spanish, German, Basque, Arabic, etc.) [26-30]. Additionally, the WordNet ontology has been utilized as an external knowledge source in various domains, such as in medical and inventive design [24, 31, 32]. The latest version of WordNet, WordNet 3.0, contains 155,287 strings and 117,659 synsets [33]. Table I shows the distribution of words across the parts of speech in WordNet.

TABLE I
DISTRIBUTION OF WORDS ACROSS THE PARTS OF SPEECH IN WORDNET

Part of Speech	Unique String	Synsets
Noun	117798	821152
Verb	11529	13767
Adjective	21479	18156
Adverb	4481	3621
Total	155287	117659

The lexical semantic representation in WordNet is very useful for natural language processing (NLP) applications, such as semantic similarity measures. Semantic similarity measures are essential to many other NLP applications, particularly word sense disambiguation, text segmentation, and information extraction [34]. In a nutshell, the semantic similarity measure represents the degree of taxonomic proximity between the concepts. The score provided by the semantic similarity measure quantifies this proximity as a function of the semantic relationship derived from knowledge sources (i.e., the WordNet ontology). Over the years, many semantic similarity measures based on the WordNet ontology have been proposed in the literature [35-40]. The measures generally fall into three categories: edge counting, information content, and feature-based approaches. Table II summarizes these approaches and their corresponding measures.

The notion underlying the edge counting approach is that the similarity between two concepts can be computed as a function of the path length that links the two concepts (i.e., the shorter the path is, the more semantically similar the concepts are) and as a function of the position of the concepts in the taxonomy. This approach views lexical ontologies as a directed graph that links concepts through taxonomic relationships, such as the is-a relationship. For instance, Wu and Palmer [35] consider the position of concepts in the taxonomy relative to the position of the most specific common concept. This approach assumes that the similarity between two concepts is a function of the path length and depth in path-based measures. The taxonomical ancestor between the terms is taken into account [i.e., the least common subsumer (LCS)] in that the measure counts the number of is-a links from each term to its LCS and also the number of is-a links from the LCS to the root of the ontology. Similarly, Leacock and Chodorow [37] also proposed a measure that considers both the number of links that connect the two concepts and the depth of the taxonomy.

The main advantage of the edge counting approach is its simplicity. The computation relies primarily on the directed graph model of a lexical ontology, which requires a low computational cost. However, because this approach considers only the shortest path between concept pairs, much of the taxonomical knowledge explicitly modeled in the ontology tends to be omitted during computation. Another known problem with this approach is the assumption that all links in the taxonomy represent a uniform distance.

TABLE II
SUMMARY OF THE EXISTING WORD SIMILARITY APPROACHES AND THE CORRESPONDING SIMILARITY MEASURES

Measure	Description	Measures
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.	• Leacock and Chodrow $sim(a,b) = -log\frac{len(a,b)}{2 \times N}$ -*len(a,b)* is the path length between a and b -N is the maximum depth in the ontology • Wu and Palmer $sim(a,b) = \frac{2 \times depth(lcs(a,b))}{len(a) + len(b) + 2 \times depth(lcs(a,b))}$ -*len(a)* and *len(b)* are the length from each term to their least common subsumer. -*lcs(a,b)* is the least common subsumer that subsumes a and b -*depth(lcs(a,b))* is the length from the root to the least common subsumer that subsumes a and b.
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.	• Resnik $sim(a,b) = IC(lcs(a,b))$ -*IC(lcs(a,b))* is the negative log of its probability occurrence. • Lin $sim(a,b) = \frac{2 \times IC(lcs(a,b))}{IC(a) + IC(b)}$ • Jiang and Conrath $sim(a,b) = IC(a) + IC(b) - 2(IC(lcs(a,b)))$
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.	• Lesk -the similarity between 2 concepts is computed from the overlapping words that exist in the corresponding glosses in WordNet

The information-content-based measure approach, on the other hand, makes use of the notion posited by information content (IC) theory by utilizing the appearance probabilities for each term in the taxonomy, which are computed based on their occurrences in a given corpus. For instance, the probability of the occurrence of a term 'x' is given in Equation 1, and the IC of 'x' is computed according to the negative log of its probability of occurrence, as shown in Equation 2.

$$p(x) = frequency(x)/N \quad (1)$$

$$IC(x) = -log\ p(x) \quad (2)$$

where N is the total number of terms that exist in the taxonomy. This measure indirectly reflects the fact that the higher the IC value is, the more specific the concept in the taxonomy is. Thus, infrequent words are considered to be more informative than common ones.

Several measures have been established using this notion, such as those of Resnik [36], Lin [38] and Jiang and Conrath [39]. Resnik [36] proposed that semantic similarity depends on the amount of shared information between two terms, which is represented by their LCS in an ontology. Two terms are considered to be maximally dissimilar if an LCS does not exist. This measure further assumes that two terms are semantically similar in proportion to the amount of information they share (i.e., the more common information the two concepts share, the more similar the terms are). Similarity measures are then based on the information content of each concept. For two given terms, similarity depends on the information content that subsumes them in the taxonomy. Lin [38], Jiang, and Conrath [39] extend Resnik's work by including the IC of both terms in the similarity computation. Lin proposed that the similarity between the two terms should be measured as the ratio of the amount of information they share and the independent information that describes the terms. The measure proposed by Jiang and Conrath [39] is based on defining the length of the taxonomical links as the difference between the IC of a concept and its LCS. This measure computes the similarity distance between two pairs by subtracting the sum of the IC of each term alone from the IC of its LCS.

Unlike the previously discussed measures, the feature-based measure is independent of the taxonomy and the subsumers of the concepts. It attempts to exploit the properties of the ontology to obtain the similarity values. It is based on the assumption that each term is described by a set of words that indicates its properties or features, such as its definitions, or 'glosses,' in WordNet. The more shared features or characteristics and the fewer non-shared features two terms have, the more similar they are. A commonly used measure utilizing this approach is the Lesk measure, which uses the

glosses in WordNet as a unique representation of the underlying terms. It computes semantic relatedness by finding and scoring overlapping features between the glosses of the two terms, as well as terms that are directly linked to them, according to the lexical ontology.

The development of semantic technology, particularly the discussed word similarity measures, provides a mechanism that enables the comparison of trademarks based on their conceptual similarity. Thus, they are also studied and incorporated in the proposed retrieval algorithm.

III. The Proposed Algorithm

The proposed retrieval algorithm is based on a conceptual model of the trademark comparison process developed in the authors' previous work [41]. It provides a bird's eye view of trademark comparison based on conceptual similarities. This paper extends the conceptual model by developing and evaluating a semantic algorithm for trademark retrieval based on conceptual similarity. The proposed algorithm employs NLP techniques and the word similarity distance method, which was derived from the WordNet ontology, together with a new trademark comparison measure. WordNet 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 [42].

In general, the retrieval algorithm consists of three main steps: the feature extraction, the hash indexing, and 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 algorithm is capable of finding similar pairs of trademarks from a database and also, in a slightly different application scenario, such as an online application, finding trademarks similar to a particular trademark. The pseudo-code that shows the steps involved in the proposed algorithm, which can be applied to the first application scenario, can be found in Appendix A.

1) Step 1: Extracting features for trademark representation in the algorithm. Each trademark is represented by two kinds of features (i.e., the trademark tokens, f_t, and the synonym list, f_s). 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. For example, the word 'flying' will be converted into 'fly.' The second feature is defined as the synonym set of the tokens and is extracted from the WordNet database. The synonym set, as defined in the context of this algorithm, includes the synonyms, the direct hypernyms, and the direct hyponyms of the corresponding tokens. Essentially, the outcome of this step yields two features: the token set and the synonym set. These are then stored to enable indexing.

2) *Step 2*: Trademark indexing using the hashing 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 f_s, i.e., the synonyms set belonging to the trademark query. The rationale for this mapping function is based on the analysis performed on the acquired infringement cases, as discussed in [41]. The final indexing table is merely a table that maps each trademark in the database to a set of trademarks from the same database for the trademark similarity computation. In this manner, the distance computation is not conducted over the entire database, which enhances the speed of the retrieval process.

3) *Step 3*: Trademark distance computation. The distance computation is based on the similarity concept introduced in Tversky's contrast theory [42]. In this theory, Tversky defines the similarity between two objects as a function of unique and shared information about the object. Based on this idea, the similarity equation between a trademark query, Q, and a trademark, T, is derived as follows:

$$sim(Q,T)=\frac{|Q_{f_t}\cap T_{f_t}|}{|Q_{f_t}\cup T_{f_t}|}+\frac{|Q_{f_s}\cap T_{f_t}|}{D}+\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}max(wordsim(x_i,y_j))}{|Q_{f_t}\setminus T_{f_t}|\cdot|T_{f_t}\setminus Q_{f_t}|}, \quad x_i\hat{I}\{Q_{f_t}\setminus T_{f_t}\}y_i\hat{I}\{T_{f_t}\setminus Q_{f_t}\} \tag{3}$$

where Q_{f_t} and Q_{f_s} are the token set and the synonyms set of the query, respectively; T_{f_t} is the token set of one of the trademarks from the database; $D=\max(|Q_{f_t}|,|T_{f_t}|)$; $Q_{f_t}\setminus T_{f_t}$ and $T_{f_t}\setminus Q_{f_t}$ are the relative complement set of T_{f_t} in Q_{f_t} and vice versa, having i and j set elements; and *wordsim* is the word similarity measure computation employed in this algorithm. In the following experiment, which aims to investigate the most suitable word similarity measure in this study, *word_sim* corresponds to the six similarity measures discussed in Table II. Fig. 1 illustrates the three steps of the algorithm, using an example from a real court case involving 'Red Bull' and 'BlueBull' as the query and the trademark from a database, respectively. In the first step, the feature extraction is performed on all trademarks in the database, including 'BlueBull.' In this step, the token and synonym sets are both extracted using the NLP and the external knowledge source, i.e., a lexicon. The mapping function indexes 'BlueBull' features in the hash table in accordance with the hashing key, in this case in the rows that correspond to the 'blue' and 'bull' keys. The trademark distance computation is then performed between the trademarks using the trademark similarity equation, as shown in Equation (3).

Fig. 2 shows an illustrative example of the trademark similarity computation between 'Red Bull' and 'BlueBull' using Equation (3). The first part of the equation is the ratio

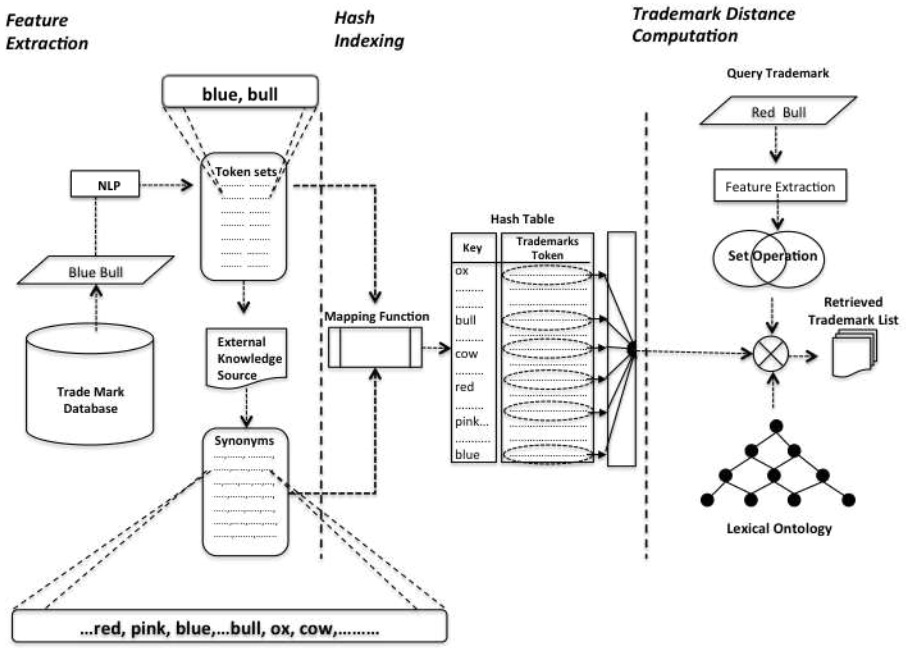

Fig. 1. An illustrative example of the steps involved for one of the trademarks from a real court case database: 'Red Bull'.

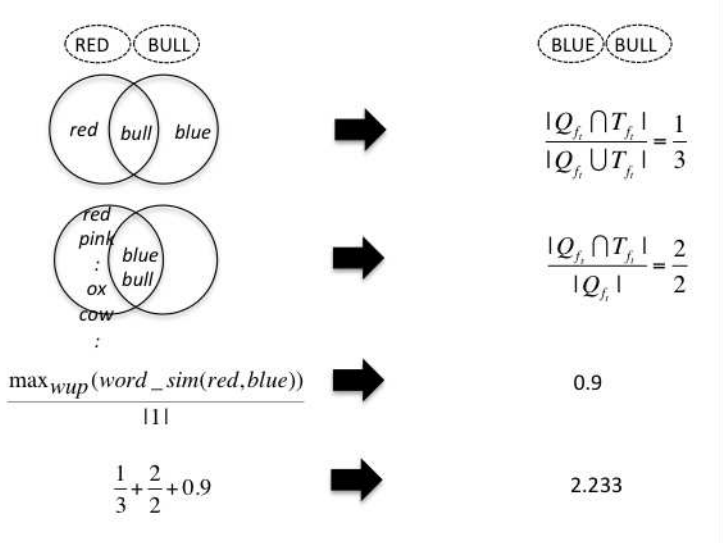

Fig. 2. An illustrative example of a trademark similarity computation using 'Red Bull' as the query and 'BlueBull' from the real court case database.

of the number of elements shared by the two trademark token sets and the number of elements in their set union operation The second part is the ratio of the number of elements in the intersection of the "Red Bull" synonyms set and "BlueBull" token set. The third part is the word similarity between the difference sets of both trademarks, measured using WordNet ontology, and the final part is the summation of the three parts, which provides the conceptual similarity score between the two trademarks.

IV. Experimental Setup

This section describes the experimental setup and the evaluation method employed in the study. A trademark retrieval system using the proposed retrieval algorithm is developed, and the algorithm is tested on two databases. Two experiments are then conducted to evaluate the performance of the proposed algorithm. The first evaluation is conducted using an information retrieval measure (i.e., R-precision score), and the second evaluation is conducted through an open call task (i.e., crowdsourcing).

A. Experiment 1

The objective of this experiment is twofold. First, the experiment examines the feasibility of the proposed algorithm against the baseline algorithm (i.e., approximate string matching). Second, it investigates the effect of employing various word similarity measures. The outcome of this study

may also suggest the most suitable word measure for use in the trademark retrieval algorithm.

In this experiment, a collection of real court cases comprised of 1,400 trademarks is obtained from [43] and analyzed as a preliminary study for the development of the retrieval algorithm. The findings from the analysis show that the cases obtained can be divided into four categories. The first category, i.e., real words, corresponds to cases involving trademark words derived from the lexical dictionary. 'Out of vocabulary' refers to trademarks with invented words, which do not have a lexical meaning. Trademarks with a combination of real and invented words are included in the 'mixture' category. The 'other' category contains trademarks with alphabetical text and family names.

Next, the analysis concentrates on the 'real words' category, which covers about 37% of the database. This category contains foreign words, words with conceptual relationships, synonyms/antonyms, and exact matching (Fig. 3). A total of 112 trademarks (see Appendix B) from 56 infringement cases that were legally proven to have conceptual similarities with earlier trademarks are extracted from this category through a manual analysis of the legal reports obtained from the disputed cases. Fig. 4 shows part of a legal report as an example. The 56 trademark pairs are then utilized as the query set to test the retrieval accuracy of the algorithm. The algorithm is tested using six word similarity measures, which are employed during the similarity comparison computation in Step 3 of the algorithm.

The R-precision score is then computed as a measure of retrieval accuracy. R-precision is a precision score at the R-th position in the retrieval result, which is also the recall score [44] in this case. The precision score is defined as in Equation 4. Because the result obtained from this experiment is the ranked retrieval result, with only one relevant trademark existing in the database for each query, the F-score, a retrieval measure normally computed for unranked retrieval results, is not a suitable indicator in this case. Hence, the precision in the first position in the retrieval for each query is computed and averaged to obtain the final score.

$$precision = |relevant\ items|/|retrieved\ items| \quad (4)$$

B. Experiment 2

The objective of this experiment is to further evaluate the performance of the proposed algorithm on a larger scale, using an open call task. The type of task is often referred to as a human intelligence task (HIT) [45-46]. Each HIT is a small portion of a large task, which is distributed among a large group of people known as workers, who have no contact with each other. The database in this experiment is comprised of 378,943 company names in the UK and Australia, which were obtained from [47]. All the entries in the database are first run as input queries, resulting in a total of six sets of 378,943 retrieval results (corresponding to the six word measures employed in the proposed algorithm). An analysis of the top retrieved results is performed to find a set of queries that produce at least three result variations from the six sets of results collected. A total of 25 queries are then randomly selected from this set. Appendix C lists the 25 queries used in the crowdsourcing evaluation and the retrieved names.

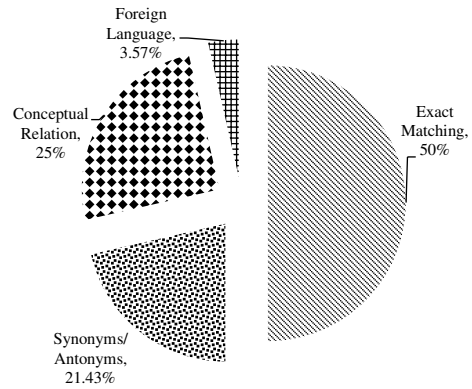

Fig. 3. Distribution of the types of conceptual similarity in the real court case database.

> The trademarks "FEEL 'N LEARN" and "SEE 'N LEARN" also ultimately suggest very similar meanings. The fact that the verbs "FEEL" and "SEE" by themselves denote different sensory perceptions does not change the fact that both trademarks contain the idea of learning with the aid of sensory organs. This fundamental idea remains in the mind of the consumer, which is why trademark similarity is also affirmed from a semantic point of view (this was also the decision of the RKGE on 21 December 2001, 3/20002172 E. 6. S. 172 – Fly away / Float away).

Fig. 4. An excerpt from the legal report obtained from one of the infringement cases.

Two crowdsourcing tasks were designed to evaluate the performance of the proposed algorithm in comparison with the traditional approximate string matching method. As in the previous experiment, the performance of the algorithm while employing various word measures is also examined.

1) Task 1

Using collective human opinions, this task compares the performance of the proposed algorithm when employing six different measures. In this task, the workers are presented with a query name and three target names. The target names are the company names extracted from the retrieval results that have the maximum similarity scores as determined by the proposed algorithm, i.e., when the six word measures mentioned above were employed. In other words, the three target names correspond to three company names returned by the proposed algorithm when using the six word measures discussed previously.

This also means that two or more results from different word measures may provide similar target names. For each of the targeted company names, workers are assigned to evaluate whether they are conceptually similar to the query names. The workers are also allowed to choose more than one targeted company name if they find them to be conceptually similar as well. This task consists of 25 HITs. For each HIT, 20 workers are assigned to complete the task. In total, 500 evaluations are obtained from this task. Fig. 5 shows one of the HITs created for this task.

Conceptual Similarity in Company Names

This task tests the existence of conceptual similarity between company names. Two or more company names may be conceptually similar if they evoke the same meaning or analogous semantic content. For example, a company with the name ***Sugarland*** may be conceptually similar to another company with the name **SWEETLAND**.

Instruction

Based on the above explanation and the company names listed below, please choose company names that are conceptually similar to the provided query. **Note: You can choose more than one company names.**

Query = PC AID

☐Pc Help Centre Ltd
☐Computer Aid
☐Pc Support Ltd

Fig. 5. HIT example for task 1 in the experiment.

Conceptual Similarity in Company Names

This task tests for the existence of conceptual similarity between company names. Two or more company names may be conceptually similar if they evoke the same meaning or analogous semantic content. For example, a company with the name ***Sugarland*** may be conceptually similar to another company with the name **SWEETLAND**.

Instruction

Base on the above explanation, please rate the conceptual similarity of the following company names.

Red Bull and **The Red Cow**

☐Highly similar
☐Similar
☐Dissimilar

Fig. 6. HIT example for task 2 in the experiment

2) Task 2

This task compares the relative performance of the proposed algorithm against the baseline algorithm, i.e., the approximate string matching algorithm, using collective human judgment as the modus operandi. The result of the proposed algorithm, when employing Wu and Palmer's word measure, is utilized in this experiment due to the findings in the previous task. In this task, the top three retrieval results from the proposed algorithm are compared to the top three retrieval results when using the approximate string matching technique. In the HIT designed for this task, workers are asked to complete a pairwise comparison in which they rate the similarity between a pair of company names (i.e., the query name and the targeted company name, which is one of the top three retrieval results). Fig. 6 shows an example of the HITs assigned in this task, in which the workers are asked to rate the similarity of the pair names from highly similar to dissimilar. Twenty workers are assigned to each query, corresponding to a total of 1,500 (25 x 3 x 20) HITs produced from the results generated using the proposed algorithm. Similar HITs are also prepared in the same manner for the retrieval results obtained when using the approximate string matching technique, resulting in a total of 3,000 HITs.

V. Result and Discussion

This section discusses the results and the performance of the algorithm used in the two experiments, together with its advantages and drawbacks from an application point of view.

A. Experiment 1: Results and discussion

Figure 7 shows the R-precision score of the proposed retrieval algorithm when employing different types of word similarity measures in the comparison computation. It also shows the accuracy of the approximate string matching algorithm, which is used in traditional text searches. It measures the capability of the algorithm to retrieve relevant trademarks in the context of conceptual similarity. All results clearly indicate that the algorithm exceeds the performance of approximate string matching by 17.6% to 20.6%. All individual results of the algorithm, when using the employed word similarity measures, surpass the R-precision score produced by the baseline algorithm. As for the performance of the algorithms when employing various word measures, the highest R-precision score is obtained when using the Lesk and Resnik measures. Both produce a score of 0.82. These measures are followed by those of Wu and Palmer, Jiang and Conrath, and Leacock and Chodorow, with a score of 0.81. The proposed algorithm produces a 0.80 R-precision score when employing the Lin measure. It can be concluded that the

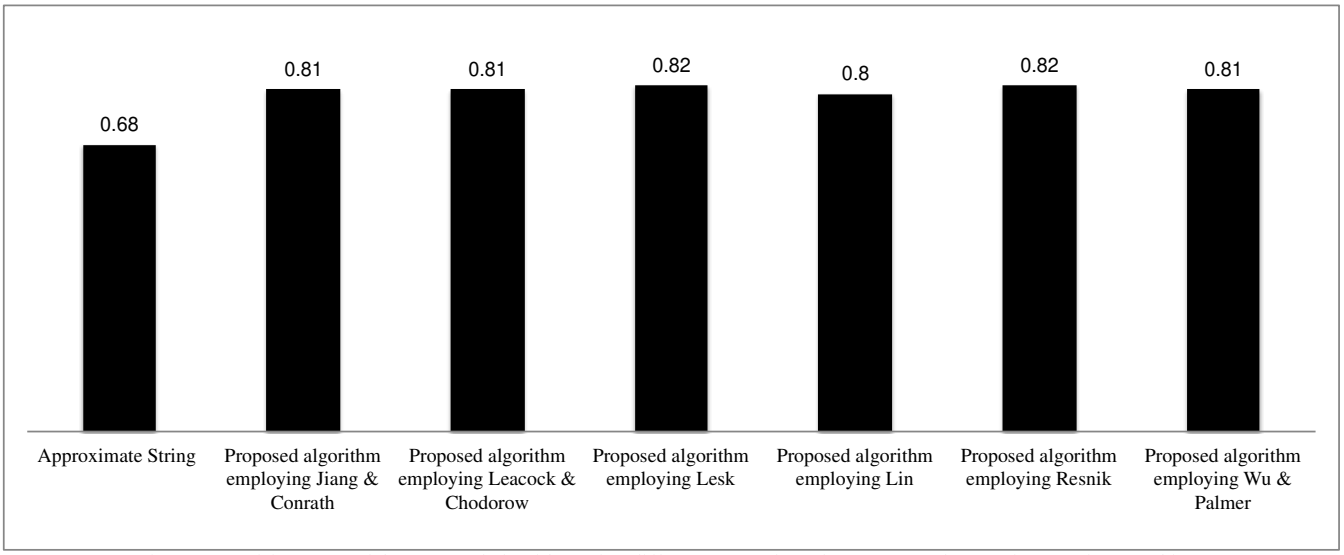

Fig. 7. R-precision score of the proposed algorithm using different types of word measures and approximate string matching.

use of various word similarity measures could affect the performance of the proposed algorithm, although the results are relatively comparable to one another. This factor is further investigated in the next experiment, using an even larger database and collective human opinion.

B. Experiment 2: Results and discussion

In the first task of the experiment, a score of 1 is assigned if the proposed algorithm retrieves conceptually similar company names, as judged by the evaluators, in every HIT. For each query, the average score from 20 workers, ranging from 0 to 1 (0 being the worst score and 1 being the best score), is computed, as shown in Table III. To analyze the results further, the average score is then divided into five scoring bands (i.e., 0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8-1). Table IV displays the results for the scoring bands, which were obtained using the six word similarity measures.

The results from the first task of the second experiment also show a similar pattern to those produced in the first experiment in that the scores vary across the table, as shown in Table IV. The results in the table also suggest that the proposed algorithm produces the highest score when using the Wu and Palmer word measure, with an average score of 0.66 (as shown at the bottom of the table). This is followed by the average scores produced using the Leacock and Chodorow and Lin measures, both scoring 0.63; the Lesk measure, scoring 0.53; and the Resnik and Jiang and Conrath measures, scoring 0.52. Likewise, the band scoring result analysis from Table IV shows that the Wu and Palmer and the Leacock and Chodorow measures produce the highest score for the band above 0.6, in which both have a cumulative count of 18 (see Table IV). However, the Wu and Palmer measure produces a slightly better score in the band above 0.8, with a count of 10. Although Lin's measure produces the highest score in the band above 0.8, with a count of 11, its total count for the band above 0.6 is 14, 16% less than the count produced by both the Wu and Palmer and the Leacock and Chodorow measures. Furthermore, the Wu and Palmer measure also produces a better R-precision score in the previous experiment as compared to the Lin measure. In general, the scores between the three measures in this section of the experiment are comparable to one another. However, because when using the Wu and Palmer measure, 72% of the results produce scores above 0.6, together with the low-complexity nature of its computation and the results from the previous experiment, this measure is considered to be a viable choice for incorporation into the proposed algorithm.

Appendix D displays the retrieval results produced by the proposed retrieval algorithm and the approximate string matching algorithm. A scoring analysis similar to that used in Task 1 is then performed, which results in the scoring shown in Table V. For each unique HIT, the average score from 20

TABLE III
THE AVERAGE SCORE FOR EACH QUERY USING THE WORD MEASURE EMPLOYED IN THIS EXPERIMENT

Queries	Jiang and Conrath	Leacock and Chodorow	Lin	Resnik	Wu and Palmer	Lesk
Red Bull	0	0.9	0.9	0.45	0.9	0.45
Imagefast	1	0	0.7	0	0	1
The Car Doctor	0.7	0.7	0.25	0.25	0.7	0.25
Landlook	0.4	0.7	0.4	0.4	0.7	1
PC AID	0.7	0.6	0.6	1	0.6	0.6
Magic Kingdom Ltd	1	0	1	0	1	0
Bodytone	0	0.95	0.95	0.1	0.95	0.1
Rug Cleaning Experts	0.7	0.75	0.5	0.75	0.75	0.5
Party Kings	0.5	0.7	0.5	0.1	0.7	0.7
Global Internet Ltd	0.35	1	0.15	1	1	1
The Letter Factory	0	0.8	0.85	0.8	0.8	0.8
Bag & Baggage Ltd	0.8	0.4	0.9	0.9	0.4	0.4
Computerman	0	1	1	1	1	1
Gas Master	0.45	0.6	0.6	0.45	1	0.6
Pet Pillow	0.25	0	1	1	0	0
Oak Tree	0.55	0.55	0.55	0.6	0.55	0.55
Sushi Kingdom	0.45	1	0.45	0	1	1
Star Ballroom	0.8	0.75	0.8	0.15	0.75	0.15
International Displays	1	0.6	0.1	0.6	0.6	0.1
Deep Sea	0.9	0.9	0.05	0.9	0.15	0.05
Planet Magazine	0.05	0.9	0.9	0.35	0.9	0.9
First Ideas	0.65	0.65	0.95	0.65	0.65	0.2
Gold Line	0.1	0.2	0.2	0.7	0.2	0.7
The Knowledge Group	0.55	0.2	0.55	0.8	0.2	0.2
The Youth Federation	1	1	0.9	0	1	1
Average Score	0.52	0.63	0.63	0.52	0.66	0.53

TABLE IV
THE AVERAGE SCORES ACROSS THE BANDS FOR EACH WORD MEASURE EMPLOYED

Scoring Band	Jiang and Conrath		Leacock and Chodorow		Lin		Resnik		Wu and Palmer		Lesk	
	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%
0<=x<0.2	6	24%	3	12%	3	12%	7	28%	3	12%	6	24%
0.2<=x<0.4	2	8%	2	8%	2	8%	2	8%	2	8%	3	12%
0.4<=x<0.6	6	24%	2	8%	6	24%	3	12%	2	8%	4	16%
0.6<=x<0.8	4	16%	9	36%	3	12%	5	20%	8	32%	4	16%
0.8<=x<=1	7	28%	9	36%	11	44%	8	32%	10	40%	8	32%

TABLE V
THE AVERAGE SCORE BETWEEN THE PROPOSED ALGORITHM AND THE APPROXIMATE STRING MATCHING ALGORITHM

Queries	Result 1		Result 2		Result 3	
	Proposed Algorithm	Approximate String	Proposed Algorithm	Approximate String	Proposed Algorithm	Approximate String
Red Bull	1.55	0.7	0.8	0.4	0.5	0.2
Imagefast	0.65	0.6	1.7	0.95	1.05	0.95
The Car Doctor	1	0.3	0.55	0.35	0.9	0.35
Landlook	1.05	0.9	0.65	0.2	0.9	0.1
PC AID	1.55	0.55	0.7	0	1.8	0.2
Magic Kingdom Ltd	1.4	0.85	0.5	0.5	1.45	0.55
Bodytone	1	0.9	1.1	1	0.9	0.9
Rug Cleaning Experts	1.45	0	1.65	0.2	1.6	1.2
Party Kings	1.1	0.65	0.75	0.75	0.65	0.6
Global Internet Ltd	1.8	1	0.85	0.75	0.5	0.5
The Letter Factory	1.15	0.2	0.6	0.3	1	0.2
Bag & Baggage Ltd	0.8	0.75	1.1	0.4	1.55	0.35
Computerman	1.65	0.95	1.9	0.9	1.55	1.2
Gas Master	1.65	1.05	0.6	0.6	1.1	0.45
Pet Pillow	0.55	0	1.5	0	0.5	0.5
Oak Tree	1.05	0.7	0.75	0	0.9	0.35
Sushi Kingdom	1.6	0.2	1.35	0.15	0.6	0
Star Ballroom	1.35	1.3	1.1	0	1.1	0.1
International Displays	1.55	0.4	0.8	0.35	0.6	0.2
Deep Sea	0.5	0.25	0.5	0.4	1.25	1
Planet Magazine	1.1	0	1.1	0.35	0.45	0.2
First Ideas	1.25	1.15	1.2	0.35	1.3	0.5
Gold Line	0.6	0.15	1	0.85	0.85	0.25
The Knowledge Group	0.75	0.7	1.45	0	0.55	0.15
The Youth Federation	1.65	0.7	1.55	0.4	0.75	0.25
Average Score	1.19	0.598	1.03	0.406	0.972	0.45

TABLE VI
THE AVERAGE SCORE ACROSS THE BANDS BETWEEN THE PROPOSED ALGORITHM AND THE APPROXIMATE STRING MATCHING ALGORITHM

Scoring Band	Result 1				Result 2				Result 3			
	Proposed Algorithm		Approx. String		Proposed Algorithm		Approx. String		Proposed Algorithm		Approx. String	
	Count	%	Count	%	Count	%	Count	%	Count	%	Count	%
0<=x<0.5	0	0%	9	36%	0	0%	17	68%	1	4%	15	60%
0.5<=x<1	6	24%	12	48%	12	48%	7	28%	13	52%	7	28%
1<=x<1.5	11	44%	4	16%	8	32%	1	4%	7	28%	3	12%
1.5<=x<=2	8	32%	0	0%	5	20%	0	0%	4	16%	0	0%

workers is computed in the range of 0 to 2 (0 being the worst score and 2 corresponding to the best score). These scores are further analyzed and grouped into four scoring bands (i.e., 0–0.5, 0.5–01.0, 1.0–1.5, and 1.5–2.0, as shown in Table VI). The analysis of the results of the second task in this experiment seeks to compare the performance of the proposed algorithm with the performance of approximate string matching as the baseline algorithm. The scores produced by the proposed algorithm exceed those generated when using the traditional approximate string matching algorithm for all 25 queries (Table V). The average score of the proposed algorithm (i.e., the scores at the bottom of Table V) for Result 1, Result 2, and Result 3 (i.e., the first three results) exceeds the approximate string matching average score by 99%, 153%, and 116%, respectively. Similarly, the results based on the band score analysis shown in Table VI further justify the applicability of the proposed algorithm because it produces much better scores than the baseline algorithm. This indirectly proves that traditional search is not suitable for a trademark search based on conceptual similarity. This type of retrieval can be performed by using the proposed algorithm, which employs a lexical knowledge source to grasp the conceptual content of trademarks.

Nevertheless, there are a few cases in which the algorithm returns conceptually irrelevant names, such as the results for the query ‘DeepSea,’ which returns ‘Seapoint’, ‘Sea Start Ltd,’ and ‘Deep Ocean Planet.’ ‘Deep Ocean Planet’ is likely to be more similar to ‘DeepSea’ than ‘Seapoint’ or ‘Sea Start Ltd’. Both ‘Seapoint’ and ‘Sea Start Ltd’ share the same token (i.e., ‘sea’), and both have an equal number of tokens (i.e., two tokens). In general, the tokens ‘deep’ and ‘point’ or ‘deep’ and ‘start’ do not seem to evoke a similar meaning in this context. However, in the lexical hierarchy, one of the senses of ‘deep,’ described as ‘the central and most intense or profound part,’ is a hyponym of ‘middle,’ defined as the ‘time between the beginning and the end of a temporal period.’ Apparently, this specific sense of the word ‘middle’ is also a hyponym of the word ‘point,’ described as ‘an instant of time.’ For these particular senses of both ‘deep’ and ‘point,’ the path length is only two nodes away. In the same way, the path length between ‘deep’ and ‘start,’ described as ‘the time at which something is supposed to begin,’ is three nodes. For this

specific part of the WordNet tree, the 'point' node is the common subsumer that subsumes 'start' and 'deep.'

In general, the shortcomings pointed out in the previous paragraph suggest that although the conceptual similarity comparison of trademarks is made possible using the proposed algorithm, this comparison is still highly dependent on the lexical ontology employed. Another point to note is that a trademark is considered to be a very short sentence. Thus, choosing the most appropriate sense of the trademark in question is highly challenging due to the limited number of words comprising the trademark. This limitation makes the common word sense disambiguation technique, which considers neighboring words, inapplicable in this context. The algorithm proposed in this paper has been tested on a database consisting of trademarks of up to seven words. Furthermore, 92% of the trademarks consist of between one to three words. The performance of the proposed algorithm has yet to be tested on longer trademarks.

The results from the experiment performed in this study also confirm that the comparison of trademarks in terms of conceptual similarity can be addressed using linguistic sources, such as a lexical ontology and lexicons. The algorithm developed in this study provides a generic mechanism for such a comparison. For example, the algorithm is not limited to the use of a specific word measure. This advantage provides a certain level of flexibility in choosing a word measure or lexical resource suited to specific applications or requirements.

VI. 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 state-of-the-art by proposing a semantic algorithm to compare trademarks in terms of conceptual similarity. The algorithm brings forward an 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.

The results from the experiment performed in this study confirm that the comparison of trademarks based on their conceptual similarity can be conducted using linguistic sources. Future work to improve the accuracy of the proposed semantic algorithm should include a study comparing the use of various lexical resources. In addition, the authors are working on extending the current approach to include retrieving trademarks with phonetic similarities and integrating their previous work on visual similarity with their new algorithms for conceptual and phonetic similarity.

Appendix A
Pseudocode Of The Proposed Retrieval Algorithm

```
Pseudocode: /*comment*/
/* This part of the code is performed for the feature
      extraction and indexing part of the algorithm*/
define f_t as the token set of a trademark;
define f_s as a set of of synonyms list that correspond
to the token set;
define f_t_all as a list of unique token extracted from the
database;
for each trademark in the database, do
  { extract f_t;
    extract f_s;
    for each token in f_t;
    { if(token does not exist in f_t_all);
        {update token into f_t_all;}}}
define hash_table as hash index table that maps token
to all trademarks in the database that contain simillar
token;
for each token in f_t_all;
  { find trademark that has similar token;
    update the hash_table;}
/*This part of code is performed during retrieval*/
for each trademark query
  { extract f_t and f_s for the query;
    map the f_s of the query to hash_table to get a list
    of trademark from the database;
    for each trademark in the extracted list from the
    hash_table
      {compute the conceptual similarity distance
    between the query and the trademark in the list}};
```

Appendix B
Trademark Pairs Extracted From The Court Cases

Trademark 1	Trademark 2
COOL WATER	AQUACOOL
Feel'n LEARN	Feel'n SEE
FRUIT TIGER	LION FRUIT
MAGIC HOUR	MAGIC TIMES
PLANE ocean	AQUA PLANET
Living Style	Lifestyle
NAVITIMER	MARITIME
PINK LADY	LADY IN ROSE
EVOLUTION	revolution
IT GIRL	It Girl
Securitas	SECURICALL
ON DEMAND	on Demand
smart home	SmartHome
NO NAME	NO NAME
THERMAL BALANCE	clima balance
FEELGOOD	FEEL GOOD
WebFOCUS	FOCUSNET
MULTI-LINE	multiline
RED BULL	BLUEBULL
GREYHOUND	greyhound
EMOTION	emotion
werkhouse	WERK HOUSE
LAWFINDER	LexFind.ch
STEPSTONE	stepping stone
SAVOUR CLUB	CLUB Saveur
Black	WHITE
SUGARLAND	SWEETLAND
tripp trapp	TRIP TRAP
COMPARIS	compare.ch
Freecom	freecom.ch
CHANEL	CHANEL
AIR FRESH	AERO FRESH
GIANTS	riesen.ch
ROYAL ELASTICS	ROYAL ELASTICS
Jetbox	JETBOXX
BULL	OX
Car4you	MOTO4YOU
BOTOX	Botoceutical
VITALITY	Vital
YELLO	YELLOW
Quiclean	fast clean
INDEX	1NDICES
MAX	MAX
Feelgood's	FEEL GOOD
MediData	medidata
DEKO LINE	DECOLINE
BIOPOINT	BIO POINT
Maxx	max
COMPARIS	comparer.ch
KICKDOWN	kickdown.ch
Bosshard	bosshard.ch
SHARK	Hai
ORPHAN EUROPE	ORPHAN INTERNATIONAL
SECRET PLEASURES	PRIVATE PLEASURES
fair assurance	fair insurance consulting

APPENDIX C

THE QUERIES AND THEIR MOST SIMILAR RETURN NAMES FOR THE SIX WORD MEASURES EMPLOYED IN THIS EXPERIMENT

Query	Jiang and Conrath	Leacock and Chodorow	Lin	Resnik	Wu and Palmer	Lesk
Red Bull	Red Cover Ltd	The Red Cow	The Red Cow	The Red Lion	The Red Cow	The Red Lion
Imagefast	Instant Image	Smart Image	Snapfast	Smart Image	Smart Image	Instant Image
The Car Doctor	Omega Car Repairs	Specialist Cars	The Car House	The Car House	Specialist Cars	The Car House
Landlook	Landcare	Land Surveys	Landcare	Landcare	Land Surveys	Property Look Ltd
PC AID	PC Help Centre Ltd	PC Support Ltd	PC Support Ltd	Computer Aid	PC Support Ltd	PC Support Ltd
Magic Kingdom Ltd	Magic City	Magic Man	Magic City	Dance Kingdom	Magic City	Magic Man
Bodytone	Mind Body Spirit	Build Tone	Build Tone	Body To Burn	Build Tone	Body To Burn
Rug Cleaning Experts	Audley Carpet Clear	Master Carpet Cleaning	Carpet-cleaning-specialist	Master Carpet Cleaning	Master Carpet Cleaning	Carpet-cleaning-specialist
Party Kings	Dancing Queen Parti	The Party Man	Dancing Queen Parties	Ace Party Co.	The Party Man	The Party Man
Global Internet Ltd	Global Network Soli	Global Web Ltd	Global Radio	Global Web Ltd	Global Web Ltd	Global Web Ltd
The Letter Factory	Mill Letter Signs	The Print Factory	The Type Factory	The Print Factory	The Print Factory	The Print Factory
Bag & Baggage Ltd	Premier Luggage &	Bag N Box	Suitcases & Bags	Suitcases & Bags	Bag N Box	Bag N Box
Computerman	Human Computer In	The Computer Guy	The Computer Guy	PC Man	The Computer Guy	The Computer Guy
Gas Master	Professional Gas Ser	Airmaster	Airmaster	Professional Gas Service	Gas Experts	Airmaster
Pet Pillow	Pets At Rest	The Pet Place	Pet Pad	Pet Pad	The Pet Place	The Pet Place
Oak Tree	The Pine Tree	The Ash Tree	The Pine Tree	Oakwood	The Ash Tree	The Ash Tree
Sushi Kingdom	The Sushi Place	Sushi World	The Sushi Place	Rock Candy Kingdom	Sushi World	Sushi World
Star Ballroom	Planet Ballroom	Star Room	Planet Ballroom	Superior Ballroom Pty	Star Room	Superior Ballroom Pty
International Displays	Global Displays	Display World Ltd	Expression International	Display World Ltd	Display World Ltd	Expression International
Deep Sea	Deep Ocean Planet	Deep Ocean Planet	Deep Red	Deep Ocean Planet	Seapoint	Deep Red
Planet Magazine	Tatler Magazine	World Magazines Ltd	World Magazines Ltd	The Daily Planet	World Magazines Ltd	World Magazines Ltd
First Ideas	An Original Idea	An Original Idea	First Concept Ltd	An Original Idea	An Original Idea	First View
Gold Line	Gold Air Internation	Goldprint	Goldprint	Silver Line Ltd	Goldprint	Silver Line Ltd
The Knowledge Group	Concept Group Ltd	Power Group Ltd	Concept Group Ltd	Knowledge Pool	Power Group Ltd	Power Group Ltd
The Youth Federation	Youth Association	Youth Association	Youth Club	Youth Service	Youth Association	Youth Association

APPENDIX D

THE THREE RETRIEVAL RESULTS FROM THE PROPOSED ALGORITHM AND THE STRING MATCHING ALGORITHM

Query	Proposed Retrieval Algorithm			Approximate String Matching Algorithm		
	Result 1	Result 2	Result 3	Result 1	Result 2	Result 3
Red Bull	The Red Cow	The Red Lion	The Red Cat	Red Bull	Red Cell	J.R Bull
Imagefast	Smart Image	Instant Image	Snapfast	Imageset	Imageware	Images
The Car Doctor	Specialist Cars	The Car House	Car Medic	The Cue Doctor	The Chair Doctor	The Tap Doctor
Landlook	Land Surveys	Landcare	Property Look Pty	Landmark	Ladbrook	Panelock
PC AID	Pc Support Ltd	Working PC	Computer Aid	P C A D	P H D	P C I
Magic Kingdom Ltd	Magic City	Magic Man	Magic World	Manor Kingdom Ltd	Gaggia Kingdom Ltd	Magic Junior Ltd
Bodytone	Build Tone	Shape and Tone	Bodytalk	Body Zone	Bodyline	Bodycote
Rug Cleaning Experts	Master Carpet Cleaning	Superstar Carpet Clean	Carpet-cleaning-specialis	Can Clothing Exports	Rendering Experts	Rgs Cleaning Ltd
Party Kings	The Party Man	Party Land	Ace Party Co.	Party Kegs	Party Link	Party Pieces
Global Internet Ltd	Global Web Ltd	Global Link	Global Radio Ltd	Power Internet Ltd	Sos Internet Ltd	Global Journey Ltd
The Letter Factory	The Print Factory	The Language Factory	The Type Factory	The Monster Factory	The Flower Factory	The Guitar Factory
Bag & Baggage Ltd	Bag N Box	Baggage Express	Suitcases & Bags	Bag & Bale Ltd	B T S Haulage Ltd	Maxi Haulage Ltd
Computerman	The Computer Guy	PC Man	Computer People	Poo Man	C M I	P C M S
Gas Master	Gas Experts	Airmaster	Professional Gas Service	Gas Matters	Car Master	G P Masters
Pet Pillow	The Pet Place	Pet Pad	Pets At Rest	Pete Hill	Pete Millson	Pet Pals
Oak Tree	The Ash Tree	The Olive Tree	The Walnut Tree	Oakmere	Fab Tec	Oakdene
Sushi Kingdom	Sushi World	The Sushi Place	Kingdoms Seafood	Cats Kingdom	Dance Kingdom	Pets Kingdom
Star Ballroom	Star Room	Superior Ballroom Pty	Planet Ballroom	Star Room	Sea Bloom	Smart Bathrooms
International Displays	Display World Ltd	Screen International	Expression International	International Diamalt	International Billiards	International Fitness
Deep Sea	Seapoint	Sea Start Ltd	Deep Ocean Planet	Deep Red	Dee Cee	Deep C
Planet Magazine	World Magazines Ltd	The Daily Planet	Magazine Creation	Piano Magazine	Flyer Magazine	Sleaze Magazine
First Ideas	An Original Idea	First View	First Impressions	First Steps	Right Ideas	Light Ideas
Gold Line	Goldprint	Silver Line Ltd	Lacegold	Fjord Line	Goldprint	Goldwins
The Knowledge Group	Power Group Ltd	Process Group	Knowledge Pool	The Knowledge Base	The Holiday Group	The Lowe Group
The Youth Federation	Youth Association	Youth Club	Youth Service	The Youth Media Ltd	The Youth Leader	NHS Support Federation

ACKNOWLEDGEMENTS

The authors wish to thank the Patent Informatics Team of the UK Intellectual Property Office for the fruitful discussions, which helped to finalize many of the ideas behind this research.

REFERENCES

[1] Office for Harmonization in the Internal Market, "Annual report 2012," [Online]. Available: https://oami.europa.eu (accessed 10.12.2013).

[2] L. Dodell, "The trademark problem: Casualty insurance's dirty little secret," 2013. [Online]. Available: http://www.carriermanagement.com (accessed 1.12.2013).

[3] U.S. Congress Joint Economic Committee, "The impact of intellectual property theft on the economy," 2012. [Online]. Available: http://www.jec.senate.gov/ (accessed 1.12.2013).

[4] C. D. Scott, "Trademark strategy in the internet age: Customer hijacking and the doctrine of initial interest confusion," *Journal of Retailing,* vol. 89, pp. 176-189, 2013.

[5] European Commission, "Report on EU customs enforcement of intellectual property rights," 2012. [Online]. Available: http://ec.europa.eu (accessed 1.12.2013).

[6] U. S. International Trade Commission, "China: Effects of intellectual property infringement and indigenous innovation policies on the U.S. economy," 2011. [Online]. Available: http://www.usitc.gov/ (accessed 20.12.2013).

[7] Office for Harmonization in the Internal Market, " 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, " 2014. [Online]. Available: https://oami.europa.eu (accessed 1.2.2014).

[8] F. Mohd Anuar, R. Setchi, and Y. K. Lai, "Trademark image retrieval using an integrated shape descriptor," *Expert Systems with Applications,* vol. 40, 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 feature matching," *Pattern Recognition,* vol. 43, 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 Recognition,* vol. 42, pp. 386-394, 2009.

[11] Y. S. Kim, W. Y. Kim, "Content-based trademark retrieval system using a visually salient feature," *Image and Vision Computing,* vol.16, pp. 931-939, 1998.

[12] J. P. Eakins, J. M. Boardman, and K. Shields, "Retrieval of trademark images by shape feature - the artisan project," in *IEEE Colloquium on Intelligent Image Databases,* 1996, pp. 9/1-9/9.

[13] J. Schietse, J. P. Eakins, and R. C. Veltkamp, "Practice and challenges in trademark image retrieval," in *6th ACM International Conference on Image and Video Retrieval*, 2007, pp. 518-524.

[14] B. Furlan, V. Batanovic, and B. Nikolic, "Semantic similarity of short texts in languages with a deficient natural language processing support," *Decision Support Systems,* vol. 55, pp. 710-719, 2013.

[15] J. Oliva, J. I. Serrano, M. D. del Castillo, and A. Iglesias, "SyMSS: A syntax-based measure for short-text semantic similarity," *Data & Knowledge Engineering,* vol. 70, pp. 390-405, 2011.

[16] P. Eagle, "News on patent, trademark and design databases on the internet," *World Patent Information,* vol. 34, pp. 326-327, 2012.

[17] Trade Mark Search. [Online]. Available: http://www.ipo.gov.uk (accessed 10.12.2013).

[18] R. Bhagdev, S. Chapman, F. Ciravegna, V. Lanfranchi, and D. Petrelli, "Hybrid search: Effectively combining keywords and semantic searches," Lecture Notes in Computer Science, vol. 5021, pp 554-568, 2008.

[19] L. Sbattella and R. Tedesco, "A novel semantic information retrieval system based on a three-level domain model," *Journal of Systems and Software,* vol. 86, pp. 1426-1452, 2013.

[20] 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 Systems with Applications,* vol. 40, pp. 2447-2461, 2013.

[21] S. A. Fadzli and R. Setchi, "Concept-based indexing of annotated images using semantic DNA," *Engineering Applications of Artificial Intelligence,* vol. 25, pp. 1644-1655, 2012.

[22] R. Setchi, Q. Tang, and I. Stankov, "Semantic-based information retrieval in support of concept design," *Advanced Engineering Informatics,* vol. 25, pp. 131-146, 2011.

[23] V. C. Storey, D. Dey, H. Ullrich, and S. Sundaresan, "An ontology-based expert system for database design," *Data & Knowledge Engineering,* vol. 28, pp. 31-46, 1998.

[24] C. Fellbaum, U. Hahn, and B. Smith, "Towards new information resources for public health - from WordNet to MedicalWordNet," *Journal of Biomedical Informatics,* vol. 39, pp. 321-332, Jun 2006.

[25] G. A. Miller, "WORDNET - A lexical database for English," *Communications of the ACM,* vol. 38, pp. 39-41,1995.

[26] E. Pociello, E. Agirre, and I. Aldezabal, "Methodology and construction of the Basque WordNet," *Language Resources and Evaluation,* vol. 45, pp. 121-142, 2011.

[27] E. Hinrichs, V. Henrich, and R. Barkey, "Using part-whole relations for automatic deduction of compound-internal relations in GermaNet," *Language Resources and Evaluation,* vol. 47, pp. 839-858, 2013.

[28] A. Fernandez-Montraveta, G. Vazquez, and C. Fellbaum, "The Spanish version of WordNet 3.0," *Text Resources and Lexical Knowledge*, vol. 8, pp. 175-182, 2008.

[29] L. Abouenour, K. Bouzoubaa, and P. Rosso, "On the evaluation and improvement of Arabic WordNet coverage and usability," *Language Resources and Evaluation,* vol. 47, pp. 891-917, 2013.

[30] J. Gonzalo, F. Verdejo, and I. Chugur, "Using EuroWordNet in a concept-based approach to cross-language text retrieval," *Applied Artificial Intelligence,* vol. 13, pp. 647-678,1999.

[31] K. C. Huang, J. Geller, M. Halper, Y. Perl, and J. C. Xu, "Using WordNet synonym substitution to enhance UMLS source integration," *Artificial Intelligence in Medicine,* vol. 46, pp. 97-109, 2009.

[32] W. Yan, C. Zanni-Merk, and F. Rousselot, "An application of semantic distance between short texts to inventive design," in *International Conference on Knowledge Engineering and Ontology Development*, 2011, pp. 261-266.

[33] WordNet Statistic. [Online]. Available: http://wordnet.princeton.edu (accessed 10.12.2013).

[34] A. Sebti and A. A. Barfroush, "A new word sense similarity measure in WordNet," in *International Multiconference on Computer Science and Information Technology*, 2008, pp. 369-373.

[35] Z. Wu and M. Palmer, "Verbs semantics and lexical selection," in *Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics*, 1994, pp. 133-138.

[36] P. Resnik, "Using information content to evaluate semantic similarity in a taxonomy," in *14th International Joint Conference on Artificial Intelligence,* 1995, pp. 448-453.

[37] C. Leacock and M. Chodorow, "Combining local context and WordNet similarity for word sense identification," in *Wordnet: An Electronic Lexical Database*, C. Fellbaum, Ed. MIT Press, 1998, pp. 265-283.

[38] D. Lin, "An information-theoretic definition of similarity," in *15th International Conference on Machine Learning*, 1998, pp. 296-304.

[39] J. J. Jiang and D. W. Conrath, "Semantic similarity based on corpus statistics and lexical taxonomy," in *International Conference on Research in Computational Linguistics*, 1997, pp. 19-33.

[40] S. Banerjee and T. Pedersen, "An adapted lesk algorithm for word sense disambiguation using WordNet," in *3rd International Conference on Computational Linguistics and Intelligent Text Processing,* 2002, pp. 136-145.

[41] F. Mohd Anuar, R. Setchi, and Y. K. Lai, "A conceptual model of trademark retrieval based on conceptual similarity," in *17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems,* 2013, pp. 450-459.

[42] A. Tversky, "Features of similarity," *Psychological Review,* vol. 84, pp. 327–352, 1977.

[43] Trade Marks Court Cases Database. [Online]. Available: http://decisions.ch (accessed 10.12.2013).

[44] C. D. Manning, P. Raghavan, and H. Schutze , *Introduction to Information Retrieval*, New York, USA, Cambridge University Press, 2008, pp. 151-175.

[45] R. Snow, B. O'Connor, D. Jurafsky, and A. Y. Ng, "Cheap and fast-but is it good?: Evaluating non-expert annotations for natural language tasks," in *Conference on Empirical Methods in Natural Language Processing*, 2008, pp. 254-263.

[46] J. R. Corney, C. Torres-, A. P. Jagadeesan, X. T. Yan, W. C. Regli, and H. Medellin, "Putting the crowd to work in a knowledge-based factory," *Advanced Engineering Informatics,* vol. 24, pp. 243-250, 2010.
[47] Company Names Database. [Online]. Available: http://uscompanydatabase.com (accessed 10.6.2013).

Fatahiyah Mohd Anuar received her B.Eng and MSc degree in Electrical and Computer Engineering from Vanderbilt University, USA, in 2000 and 2002. In 2015 she completed her PhD studies in Engineering at Cardiff University, UK.

From 2002 to 2007 she worked as an engineer in a telecommunication company in Malaysia. She then worked as a researcher/lecturer in Multimedia University, Malaysia. Her research interests include image processing, semantic computing, affective computing and data mining.

Rossi Setchi received her degree in Mechanical and Electrical Engineering from Moscow Technological University in Russia, and her PhD from Cardiff University, UK. She is a Chartered Engineer, Chartered IT Professional, European Engineer and Fellow of the Institution of Mechanical Engineers, Fellow of the Institution of Engineering and Technology, Fellow of the British Computer Society, and Senior Member of IEEE. She acts as an Associate Editor of the International Journal of Systems Science (IJSS), and the International Journal of Knowledge and Systems Science (IJKSS).

Professor Setchi leads the Mechanics, Materials and Advanced Manufacturing Research Theme which currently comprises over 30 academics and 50 researchers and PhD students. She also leads her Knowledge Engineering Systems (KES) research group. Her main research interests are in intelligent manufacturing, robotics, systems engineering, AI, knowledge-based systems, and knowledge engineering.

Professor Setchi has been a recipient of the 1999 Literati Club Award for Excellence, 2004 Donald Julius Groen Prize of the Institution of Mechanical Engineers, and awarded the KES Conference Best Paper Award in two consecutive years, 2012 and 2013.

Yu-Kun Lai Yu-Kun Lai received his Bachelor's and PhD degrees in Computer Science from Tsinghua University in 2003 and 2008, respectively. He is currently a Lecturer of visual computing in the School of Computer Science & Informatics, Cardiff University, UK. His research interests include computer graphics, geometry processing, computer vision and information retrieval. He is on the editorial board of The Visual Computer.