Building an Internet-Based Knowledge Ontology for Trademark Protection

Charles V. Trappey, National Chiao Tung University, Taiwan

 https://orcid.org/0000-0002-3069-6702

Ai-Che Chang, Shih Hsin University, Taiwan

 https://orcid.org/0000-0002-1524-5225

Amy J. C. Trappey, National Tsing Hua University, Taiwan

 https://orcid.org/0000-0001-7651-7012

ABSTRACT

Global online sales for products, where many are substantially identical or deceptively similar, are the cause of a growing number of trademark (TM) infringement lawsuits. This research proposes an intelligent trademark legal precedent recommendation system to assist trademark owners to find relevant past cases, laws, and judgments to form legal arguments to defend against infringement. Judicial precedent and applicable laws from the USA are used to construct an ontology of trademark litigation knowledge. The ontology is used to analyze potential infringement cases with similar laws and precedents used to resolve previous legal disputes. The analysis provides a basis for proceeding with legal action necessary to protect a company's brand equity when arguing potential trademark infringement. Using the Python programming language, the precedent-based recommendation system provides a means for continuously updating trademark case data and assists TM owners to quickly identify similar cases to support infringement allegations.

KEYWORDS

E-Discovery, Infringement Analysis, Ontology, Recommendation System, Text Mining, Trademark Protection

1. INTRODUCTION

More than 3,000 trademark infringement lawsuits are filed each year in US district courts with internet commerce and grey channels (Ertekin, Sorescu, & Houston, 2018). Unscrupulous and lawless retailers and wholesalers infringe upon existing trademarks and exploit the anonymity and lack of regulation of the internet and import regimes. The goal of online trademark protection is of great importance since the uniqueness of a business or a brand is expressed through the establishment of a trademark and represents significant intangible value (Howard, Kerin, & Gengler, 2000). Intangible assets such as brands and trademarks yield extremely profitable and life-long earnings to enterprises. Companies are increasingly focusing on building their own trademarks and brand stories to gain market advantage

that differentiates their offerings from other companies. Corporations that invest in brand equity can register and use global laws to protect against infringement.

Mobile devices provide various applications and multiple channels for online advertising in the digital world. Advertisers that are intellectual property (IP) infringers convey misleading brand stories, product information, and advertisements to consumers. The IP infringement attacks use digital media graphics, images, songs, messaging and even games. Digital advertising expenditures worldwide have exceeded US 283 billion dollars in 2018 and are predicted to surpass US 333 billion dollars by 2019. For the first time, digital advertising will account for roughly half of the global market advertising budget (Enberg, 2019). Digital advertising fraud is predicted to cost advertisers between US 23 to 30 billion dollars in 2019, representing up to 9% of total digital advertising expenditures (Cavazo, 2019). During the year 2015, for every US three dollars spent on digital advertising, fraud accounted for one US dollar (Adage.com, 2015). Brand equity losses to counterfeits are estimated to be US 250 billion dollars in the US each year and as much as US 1 trillion dollars globally (Ertekin et al., 2018). Low-cost competition with poor quality and fake merchandise causes original trademark owners to lose loyal customers that are gained through good will, augmented services, and significant financial investment to promote name recognition and market position. One reason for rampant infringement is that trademark holders are overburdened with litigation that advances to trials costing between US 375 thousand to US 2 million dollars per case (Ertekin et al., 2018). Surveillance of global digital content published on websites are difficult to trace and identify the identities and country of origin of the IP infringers. Recommender systems are an effective and fast sampling method which may be used to produce a lead on infringers of IP (Hamidi & Mousavi, 2018).

For protecting trademarks and avoiding infringing on the rights of TM owners, the study provides an intelligent precedent-based trademark litigation recommender system to assist trademark owners to find relevant past cases, laws, and judgments to form legal arguments to defend against infringement. As an initial effort to protect the brand equity of trademark owners, this research proposes a methodology to identify potential trademark infringement and aggregates empirical evidence against violators. The construction of the ontology creates relationships between the knowledge map of trademark infringements using a corpus of judicial case law precedents. The information of precedents is organized using text mining of the litigation procedures to transform qualitative information into quantitative measures. The information provides a source of data that can be continuously updated and provides a legal infringement corpus database. The outcome of this research provides trademark owners with past legal settlements and trail litigation outcomes that have established precedent for supporting the decision-making process of determining future case judgments. This research paper is organized as follows. Section 2 reviews and discusses the background literature. The proposed three phase precedent analysis and benchmark methodologies are described in Section 3. Section 4 presents the case implementation. The concluding remarks for implementation are provided in Section 5.

2. BACKGROUND AND RELATED WORK

In order to understand the context of the article, the study presents literature related to the definition of trademark and trademark infringement, trademark protection laws, and judicial precedent text mining. The definitions form the basis for the research theme and methodologies.

2.1. The Definitions of Trademark and Trademark Infringement

According to the United States Patent and Trademark Office (USPTO), a trademark is a word, phrase, symbol, or design, or a combination of these elements, that identifies and distinguishes the source of one party's goods from those of others. The terms trademark or mark refer to both trademarks and service marks which are inscribed on products, related to products, and are used by numerous media communications including the internet, radio, or television. Trademarks as brand names provide measurable financial equity to enterprise products or services. The trademark provides a unique,

measurable asset that creates a meaningful relationship with customers. The trademarked brand name differentiates the companies offering from other products and provides a platform from which the company can build a competitive and sustainable advantage over competitors (Keller, 1998). The customer's confidence and loyalty depend upon the trust, belief, and augmented services that are represented by the trademark. Cohen (1995) notes that a trademark does not protect a product; it protects a product's identification. For protecting commercial profits and establishing business rights, registering a trademark is the safest way to maintain exclusive rights to a company's proprietary sign. Trademark owners must complete their renewal applications and pay trademark fees to the USPTO for each 10-year period of registration (USPTO, 2018). A consumer buying a brand is guaranteed to receive the benefits and services described by the legal agreement provided by brand owners to their customers.

The global online media has significantly contributed to the establishment of world brands. At the same time, global media, such as the internet and internet sales sites, have led to increasing cases of trademark infringement and brand confusion that are difficult to identify and prosecute. Companies spend millions of US dollars to establish global brands as can be seen by the amounts of money spent for commercials played during the US National Football League playoffs every year (Statista, 2018). Trademarks that are infringed upon lose brand equity, decrease commercial benefits, and sometimes irrevocably damage the brand and trust held by customers (Scoot, 2013). In order to avoid the likelihood of confusion and misrepresentation, the purpose of trademark law is designed to prevent unfair competition by protecting the use of identifiers (a symbol, word, logo, slogan, design, or domain name) that uniquely distinguishes the genuine goods or services of a firm (Howard et al., 2000).

The Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS Agreement) is the most comprehensive and influential international treaty on intellectual property rights. Article 16 of the TRIPS Agreement stipulates that a registered trademark owner shall have an exclusive right to prevent others from obtaining consent to use their exclusively owned intellectual property protecting their brand or trademark (WIPO, 1996). The use of the same or similar products or services with the same or similar identification during commercial transactions may result in the likelihood of confusion. Any use of the same identifier for the same product or service may be misleading and prejudicial to the brand owner. The above rights shall not infringe upon any existing rights, and shall not affect the possibility of members granting rights based on their use (WTO, 1994). Thus, the likelihood of confusion is the essential criteria used to determine trademark infringement (Allen, 1991).

2.2. Trademark Protection Laws

The Federal Trademark Law codified in the Lanham Act (2000) (1) defines the legal definition of trademark infringement, (2) clarifies the remedies that trademark owners can seek, such as injunctions and civil compensation, (3) establishes criminal penalties for selling counterfeit goods, (4) specifies the requirements of trademark registration, and (5) formulates the specific rights for trademark owners. Given the above trademark law (Figure 1), most trademark judgments determine whether the infringement is consistent with prior rulings of the US Federal Court (Scott, 2013). First, the legal dispute settlement procedure determines if the trademarks are used in commerce within the scope of the underlying purposes of trademark law: market efficiency and market efficacy. In the case of commercial use, customer confusion should be considered by the strength of plaintiff's trademark, the similarity of the trademark, the relevance of the product, the relevance of the marketing channels, the level of customer care, actual confusion evidence, and the defendant's intention (Albert et al., 1999).

Initial interest confusion is judged if the commodities or advertisements do not involve customer confusion. Trademark infringement occurs when one party uses a mark that is identical or similar to another party's mark, creating a likelihood that consumers will be confused by the source or origin of the goods or services which deceptively represents the wrong merchant (McCarthy, 1998). Trademark law allows a finding of infringement if a consumer knows the actual source or origin of the goods or

Figure 1. Federal court analyses of internet trademark cases
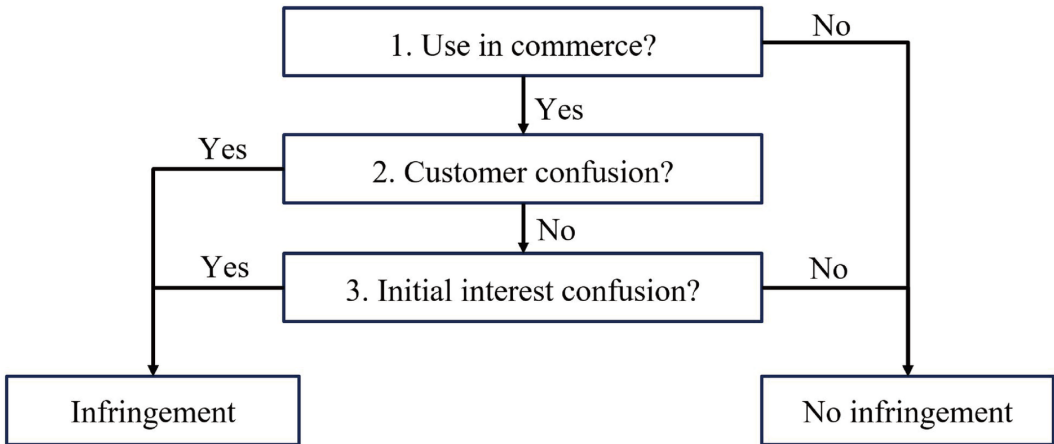
services at the time of purchase and is not confused and knowingly buys a fake product (Scott, 2013). Even if consumers make an erroneous or a deliberate purchase, trademark infringers are causing confusion and are diverting potential consumers away from the original source of product or service while wrongfully profiting from the goodwill of the registered brand. According to U.S. Trademark Law (USPTO, 2017), Figure 1 depicts the evaluation procedures and the relevant regulations are shown in Table 1.

2.3. Judicial Precedents Text Mining

Text mining is a structured method used to automatically detect trends, relationships, and to extract patterns of knowledge from unstructured text (Hearst, 1999; Tan, 1999). The method is viewed as an extension of knowledge discovery or data mining (Simoudis, 1996; Raguseo, Pigni, & Piccoli, 2018;). Fan et al. (2006) note that text mining has been applied to business management, health care, government services, education, and manufacturing. Social networks, microblogs, and e-commerce platforms provide large volumes of data that can be extracted and analyzed using text mining techniques (Yuce et al., 2014; Han et al., 2018; Trappey et al., 2016a; Trappey et al., 2016b). These studies provide fundamental text mining methods used to extract information for decision making and monitoring of IP status. There are several components of text mining including text segmentation, summary extraction, stemming, keyword identification, topic detection, taxonomy generation, term

Table 1. The U.S. trademark laws (USPTO, 2017)

Law (USC*)	The Scope of the Law
15 USC§1114	Remedies; infringement; innocent infringement by printers and publishers
15 USC§1115	Registration as evidence of right to exclusive use; defenses
15 USC§1117	Recovery of profits, damages, and costs
15 USC§1125	False designations of origin, false descriptions, and dilution forbidden
15 USC§1127	Construction and definitions; intent of chapter
28 USC§1291	Final decisions of district courts

*The Code of Laws of the United States of America (abbreviated to U.S. Code or USC) is the official compilation and codification of the general and permanent federal statutes of the United States. The main edition is published every six years by the Office of the Law Revision Counsel of the House of Representatives, Washington, DC.

clustering and document categorization. With these components, unstructured text documents are transformed into knowledge whereby decisions can be made without bias (Berry, 2004).

In order to identify key terms explicitly, input data is organized and stop words (words with little meaning) are removed. Words stems are created by removing the affixes and suffixes to derive the most meaningful root words for processing. When extracting the terms from documents, the term frequency (TF) and inverse document frequency (IDF) are used. TF calculates the occurrence frequency of a specific term in one document. Considering document length, the TF is normalized to increase the accuracy of term usage. IDF measures the rarity of a term across many documents (Trappey et al., 2010). Salton and Buckley (1988) combine TF and IDF (TF-IDF) to represent the representative term which appears most frequently in the document but rarely across a collection of documents. This research uses normalized term frequency and inverse document frequency (NTF-IDF) to weight each term for extracting key features from the trademark litigation and cited precedents of trademark infringement cases. Key term extraction creates a list of key terms related to these precedents and the list may be updated as new information is provided online.
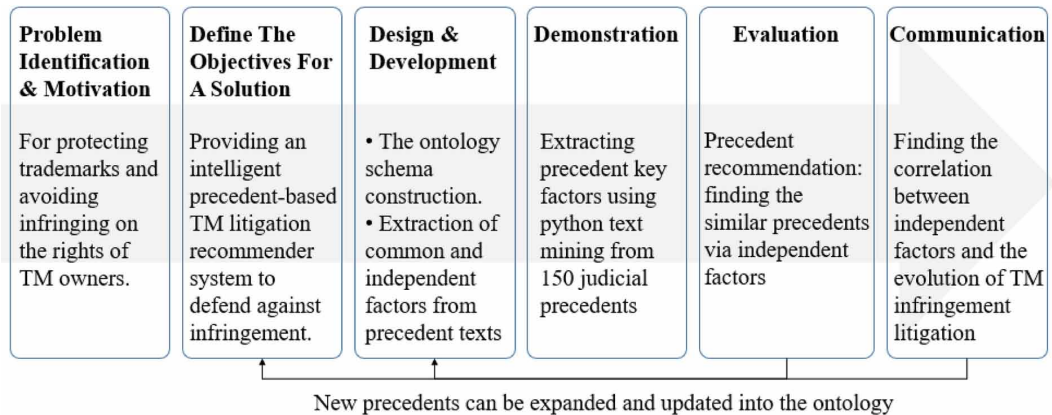
3. METHODOLOGY

The research proposes an intelligent precedent-based trademark litigation recommender system using an ontology constructed from published US court legal precedents and judicial outcomes.

According to design science in information system research methodology (Peffers et al., 2007; Hevner et al, 2004), the study constructs an infringement ontology of the precedents and develops a legal precedent analysis flowchart to identify the established process flows during litigation (Figure 2). The process model is divided into six activities in sequence: problem identification and motivation, define the objectives for a solution, design and development, demonstration, evaluation, and communication. Problem identification and motivation and define the objectives for a solution have stated in the above two sections. In this section, the methodology is divided into three main phases: formulating the criteria and processes for identifying case precedents, the ontology schema construction, and extraction of common and independent factors from the documented precedent texts. Every field creates ontologies to limit complexity and organize information into data and knowledge. As new ontologies are made, their use hopefully improves problem solving within that domain. This study constructs an ontology schema of TM litigation knowledge for understanding the overall structure of the TM litigation precedent. A computer assisted method incorporating text mining techniques is the core methodology used to extract the common and independent factors. These factors are independently judged by a trademark expert. The utility of this methodology makes the documented TM litigation precedent be the machine-readable functionality. Formulating the criteria, processes, and the common and independent factors for the TM litigation precedent fulfills the extraction function of the computer-assisted system automatically. One benefit of the machine-readable functionality facilitates the outcome of the precedent-based recommendation system with efficacy and consistency. The follow-up phases (i.e., demonstration, evaluation, and communication) analyze the actual judicial precedents in the next section. The study builds the intelligent system using the Python programming language. The research results measure the correlation between each precedent based on the mining results. High repetition of key words between each precedent represents the high correlation between infringement case precedents.

3.1. Formulating the Criteria and Processes for Precedents

The research establishes an intelligent trademark legal case recommender system to judge whether a case includes an act of infringement. The criteria and process used follow the standards of US legal statutes and the relevant laws. The process of precedent analysis follows.

Figure 2. A legal precedent analysis flowchart based on the design science in information system research methodology

Problem Identification & Motivation	Define The Objectives For A Solution	Design & Development	Demonstration	Evaluation	Communication
For protecting trademarks and avoiding infringing on the rights of TM owners.	Providing an intelligent precedent-based TM litigation recommender system to defend against infringement.	• The ontology schema construction. • Extraction of common and independent factors from precedent texts	Extracting precedent key factors using python text mining from 150 judicial precedents	Precedent recommendation: finding the similar precedents via independent factors	Finding the correlation between independent factors and the evolution of TM infringement litigation

New precedents can be expanded and updated into the ontology

3.1.1. The TM Infringement Precedent Analysis Process

The relevant TM infringement precedents are used as references to establish the legal corpus for the recommendation system. The selected precedents include information about the participants (plaintiffs and defendants), industry, region, cause of litigation, plaintiff's claim, judgment result, and applicable laws. The ontology knowledge map of the precedents is constructed manually while analyzing the TM infringement precedents (Figure 3) and are verified by the domain expert. The outcome of the process is based on the judgment result. If the judgment results in conviction, then the laws are extracted to document and categorize their usage in the precedents.

3.1.2. The Method and Process of Precedent Classification

The analysis and classification of the infringement precedent depends upon the judicial decision and the associated factors. The precedent is divided into sections including name, number, participant (plaintiff, defendant), industry, date, cause, agreement, and result. The classification method helps identify similar precedents and refer litigation results with similar factors of classification.

Figure 3. The analysis process based on TM infringement precedents
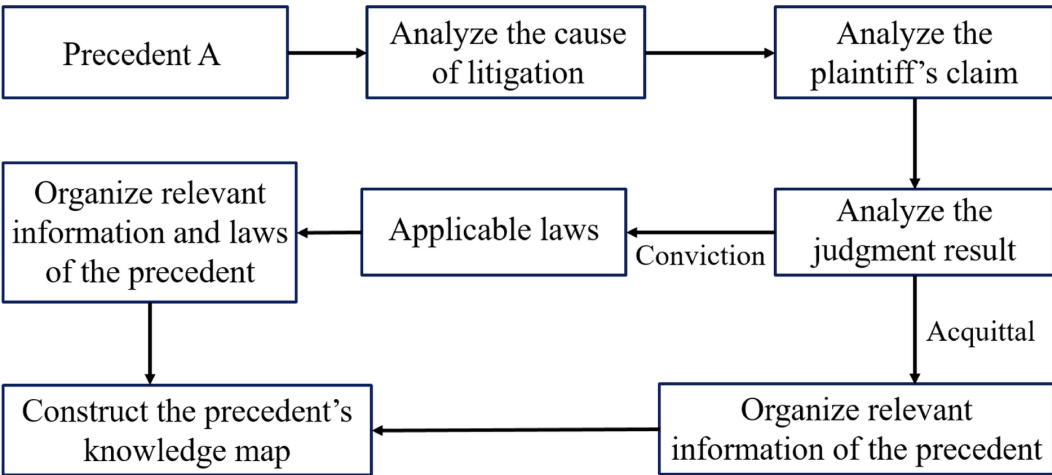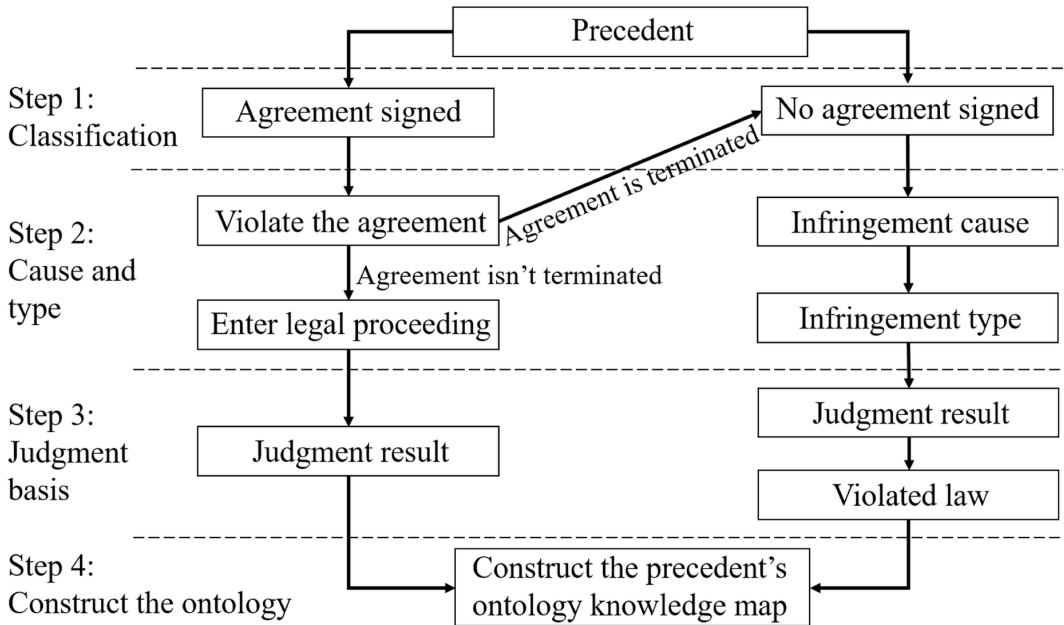
Figure 4. The method and process of precedent classification

The signing of an agreement between the two parties differentiates trademark infringement precedents into two types (Figure 4). This step has a critical impact on the judgment result. The second step is to clarify whether the signed agreement was breached. Given the period of the agreement, the research documents the judgment results and conditions. For non-contractual infringement, the cause of and type of infringement and the judgment is recorded. The infringement ontology map may be continuously expanded, updated and improved with the inclusion of new precedents.

3.2. The Ontology Schema Construction

Constructing the trademark service platform uses the Protégé software to define and organize the litigation ontology knowledge. An ontology is an organizational system designed to categorize and explain relationships between concepts of knowledge in a specific domain (Swartout & Tate, 1999). The construction of the precedent ontology and the introduction of WebProtégé are provided in this section.

3.2.1. The Construction Logic of a Precedent Ontology

After analyzing precedents and the related laws used to argue outcomes, the data and factors are extracted to establish the map of the TM infringement precedents. The components of the ontology include classes, attributes, and instances. The circles represent the class, the words on the black line represent the instance, and the boxes represents the instance's attribute. In Figure 5, a precedent is draw as the starting point of the map from which several sub-classes, entities, TM infringement types, judgment types, and applicable laws are accounted for. In the sample precedent, Structures v J.P., Structures is an instance and this instance has different attributes such as a no agreement status, a judgment result, and a type of conviction. As new precedents are added, more detailed sub-classes, instances, and attributes are generated to construct a more complete TM infringement ontology.

Figure 5. The initial structure of a TM precedent map
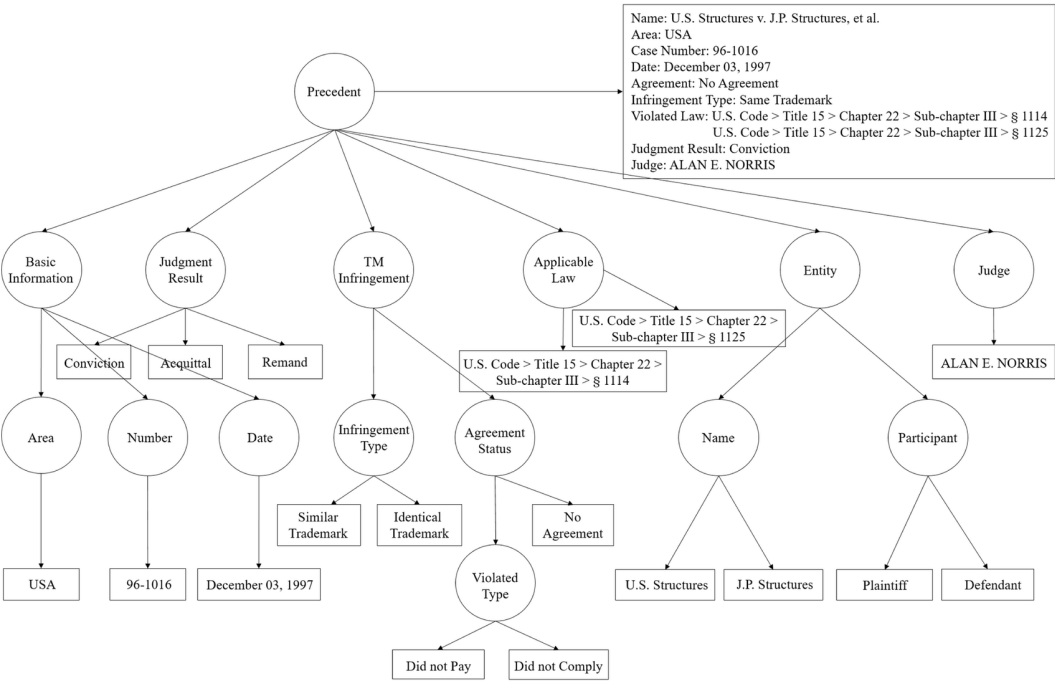
3.2.2. The Introduction of WebProtégé

WebProtégé software was developed by Stanford University in 2016 and is commonly used in the construction of ontologies. The original Protégé system used Java and was offered to the public as open source software. The flexible design and numerous plug-in modules helped launch the software as one of the most widely used knowledge-based ontology building software systems in the world. A notable feature of WebProtégé is that it allows users to edit online while sharing knowledge among users. WebProtégé ontologies can be expanded continuously based on user contributions and the addition of new knowledge. The technology and laws of TM infringement change over time. Since refining and sharing infringement knowledge is needed, WebProtégé was selected as the tool to construct the ontology for this research.

3.2.3. The Construction of the Initial TM Litigation Ontology

The research proposes a method to construct an ontology of TM litigation knowledge. The collected precedents are integrated into a precedent class which organizes the factors of the class, instances, and attributes. The infringement ontology is shown in Figure 6. Understanding the infringement type and agreement status of new and existing cases enables the user to search for similar or corresponding infringement judicial precedents in the system and to identify similarities with precedents.

3.3. Extraction of Common and Independent Factors From Precedent Texts

The research to date has identified eleven essential classes or subclasses. Text mining is used to simplify complicated judicial precedents to representative parts which can be quickly retrieved and reviewed for similarity and referencing. The eleven classes are divided into common and independent factors and are defined in this section. The common factors are the basic facts of each precedent which have a fixed format while the independent factors have unique characteristics which vary with the wording of the legal judgments.

Figure 6. The ontology of trademark infringement (brief overview)
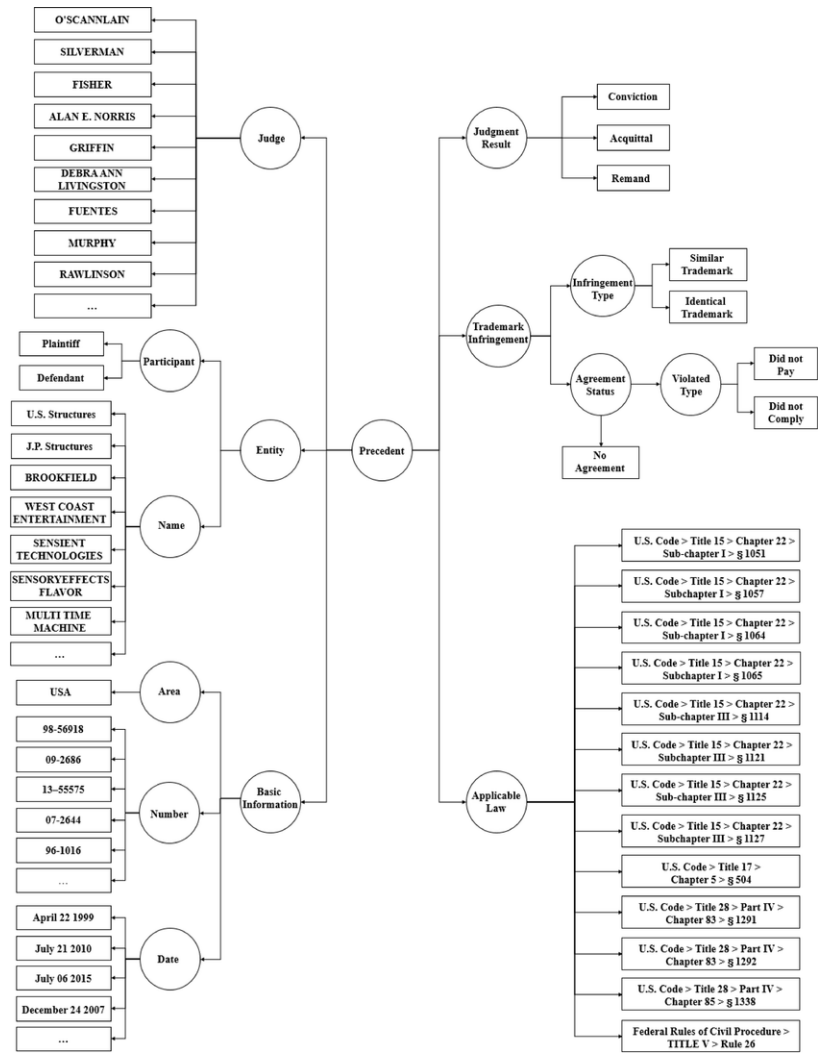
3.3.1. Retrieving the Common Factors From Precedent Texts

The common factors include the name, number, date, area, plaintiff, defendant, and judge. Using the proposed computer supported techniques, these factors are automatically retrieved from the precedents.

Name

Plaintiffs and defendants appear in the second paragraph of the title. The sample is shown in Figure 7. The research mines the second paragraph of the heading and sets "v." as the dividing point between the plaintiff and the defendant. The legal precedent name is extracted from the first word in the second paragraph and the first word next to "v." The expression "et al." is used to represent multiple parties. The final output of the precedent name is shown as "U.S. STRUCTURES v. J.P. STRUCTURES, et al."

Figure 7. The heading format of the judicial precedent
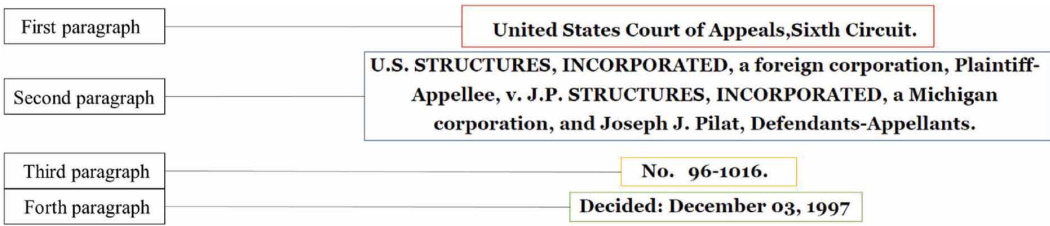
Number

Under the header format of the precedent in Figure 7, each precedent has a unique number located in the third paragraph. After removing some stop words (e.g., no. and number), the precedent number is extracted using text mining. The final output of the precedent number is shown as "96-1016."

Date

Each precedent has a fixed format judgment date which is located in the fourth paragraph of the heading. The text mining program deletes stop words so as to include only the date. The final output of the judgment date is "December 03, 1997."

Area

The legal area of the judgment is most frequently shown in the first paragraph of the heading. The text mining program extracts the full name and acronym of each state in USA. Second, the study mines the first paragraph of the heading and checks whether the defined state names or words are retrieved. Third, the precedent is classified into a USA state judgment or a regional judgment.

Plaintiff and Defendant

The participants (the plaintiffs and the defendants) are located in the second paragraph of the heading. The text mining program uses the heading and adds "v." as the separation point between the plaintiffs and the defendants.

Judge

The presiding judge should carefully analyze the precedents since judges often have specialties and unique perspectives of analysis. The data preprocessing identifying the judge is found at the end of the precedent text where the judge's signature is located.

3.3.2. Mining the Independent Factors From Precedent Texts

Each judicial precedent is evaluated by a judge using their independent logic and writing style which generates unique characteristics and conclusions for each case which is experienced based. The outcomes are seldom biased or they would be appealed. However, each judge's unique experience contributes new knowledge to the analysis. The research formulates a judgment prediction for the expected outcome before applying text mining. Using the constructed ontology of trademark

Figure 8. The decision process of agreement between the plaintiff and defendant
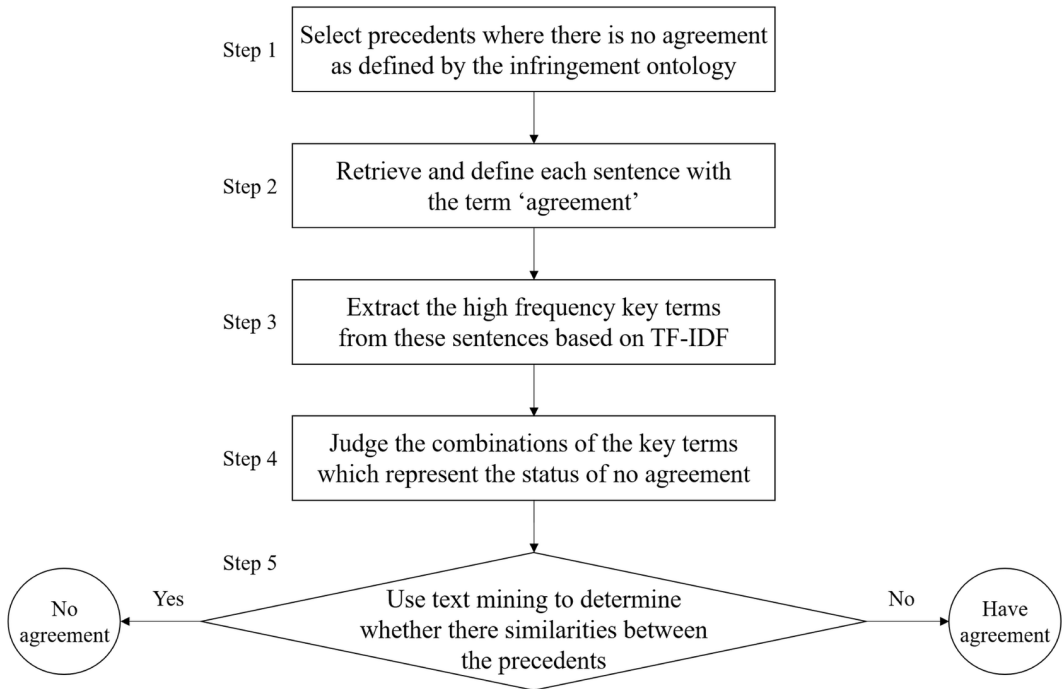
infringement, the independent features of precedents are divided into agreement status, judgment result, infringement type, and applicable law.

Agreement Status

Trademark infringement relates to whether an agreement or relevant authorization is signed between the infringer and the trademark owner. Criminal motive is different from unauthorized use. The research decision process is illustrated in Figure 8. First, select precedents where there is no agreement as defined by the infringement ontology. Second, retrieve and define each sentence with the term 'agreement.' Third, extract the high frequency key terms from these sentences based on TF-IDF algorithm. Fourth, remove irrelevant key terms and judgments where the combinations of these key terms represent the status of no agreement (e.g., terminate vs. agreement, breach vs. agreement, and violation vs. agreement). Fifth, text mining is used to determine whether there are similarities between the precedents.

Judgment Result

This study is based upon the logic defined by the US federal court proceedings in Figure 9. The litigation proceedings begin with the Federal District Court. The second instance enters the United States Courts of Appeals, and finally the Supreme Court of the United States. In the proceeding, the plaintiff or the defendants may appeal to a higher court if there are relevant concerns with the judgment result. According to the judgment result of the higher court, parts of the judgment result may be returned to the lower court. Once both sides have no objection to the judicial system rules, the litigation is terminated.

Figure 9. The US federal court proceedings
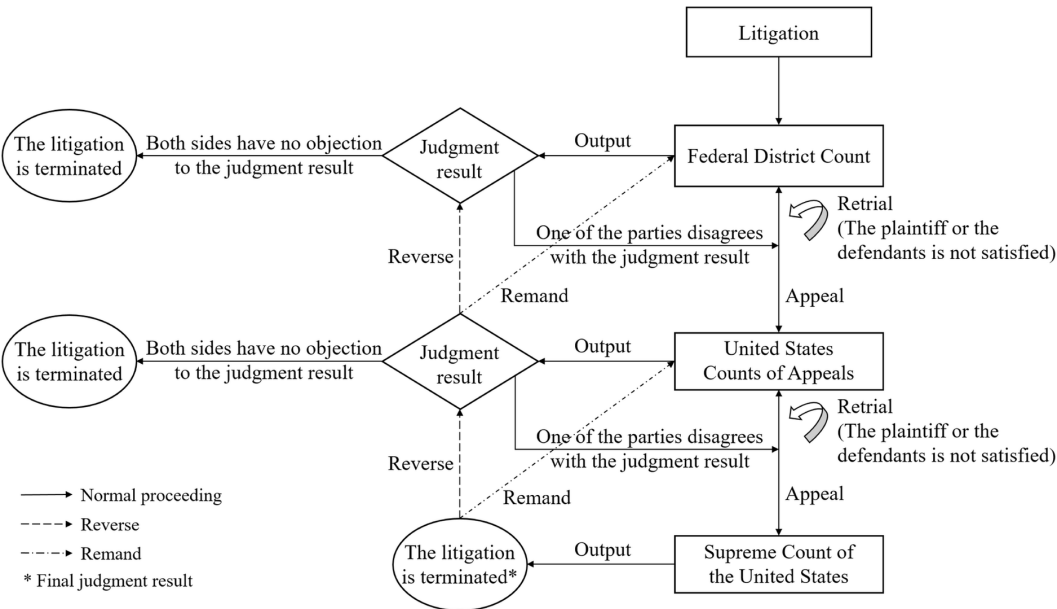
Many precedents enter a second trial at the United States Courts of Appeals. This research classifies precedents into three types based on the status of the appellants and the flow depicted in Figure 10.

Type A: If the appellant is the defendant, the research infers that the defendant is guilty in the result of the first instance.
Type B: If the appellant is the plaintiff, the research infers that the defendant is no guilty in the result of the first instance.
Type C: If the appellant is the plaintiff, the research infers that the defendant is guilty but the plaintiff is dissatisfied in the result of the first instance.

For type A, the research infers that the previous judgment is judged as a conviction while the appellant is the defendant. Using text mining, the research verifies whether the current judge agrees or disagrees with previous judgment results. If there is agreement, the judgment result is a conviction. Otherwise the outcome is an acquittal or remand.

Regarding Type B and Type C cases, the research uses text mining to evaluate the conviction or acquittal for the defendant in the previous judgment result. Under the status of acquittal, if the second trial judge holds a consistent result with the first trial, the judgment result is an acquittal. Otherwise, the case law reverses the remand for the problematic part. Under the status of conviction, if the current judge provides consistent results with the previous trial, the judgment will result in a conviction or remands. Please note that if the current judge has inconsistent results with previous trials, the computer support approach will find the discrepancies.

Infringement Type

Using the infringement ontology, the infringement type is divided into deceptively similar trademarks and identical trademarks. The original meaning of a deceptively similar trademark is that

Figure 10. The determination process of judgment results
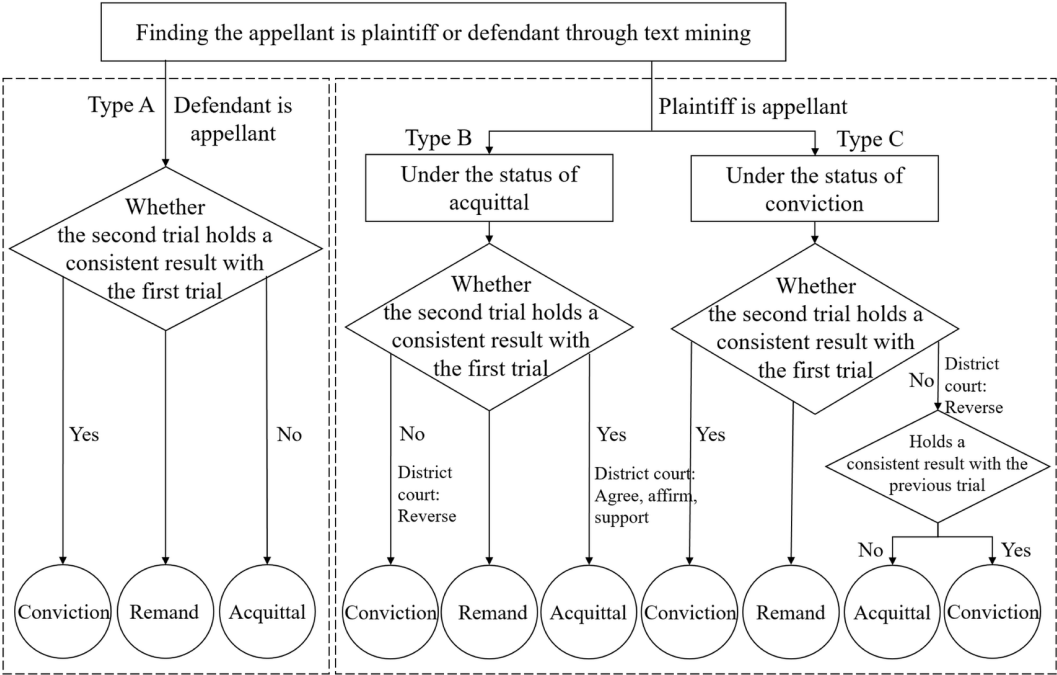
the trademark causes customer confusion. The research defines eight items for customer confusion based on 15 USC§1114(1) and 1125(a). If the proposed items or the composed similar terms are text mined, the precedent is identified as a deceptively similar trademark and the disputed items are listed:

1. Strength or Weakness of the Plaintiff's Mark
2. Defendant's Use of the Mark
3. Similarity of Plaintiff's and Defendant's Marks
4. Actual Confusion
5. Defendant's Intent
6. Marketing/Advertising Channels
7. Consumer's Degree of Care
8. Product Line Expansion

Either the trademark is deceptively similar or identical. The proposed research uses TF-IDF to highlight high frequency key terms which appear in the related legal documents. Violations of trademark precedents are provided as examples to demonstrate the efficiency of text mining. The extracted key terms are logged in the database as used by the infringer under violation of the cooperation agreement. Table 2 provides a sample of combined sentences containing high frequency root words. A total of thirteen combinations match the criteria of deceptively similar trademarks violating the agreement between the signing parties.

Applicable Law

The quoted laws are presented in a standard format for the legal case precedents. The computer automatically retrieves the cited laws which are relevant to the precedent. All applicable laws can

Table 2. Root word combinations belonging to disputed trademarks

1. …'breach*'… 'agreement'…
2. …'breach*'… 'contract'…
3. …'violat*' …'contract'…
4. …'violat*' …'agreement'…
5. …'breach*'… 'franchi'… 'contract'…
6. …'violat*'… 'franchi'… 'agreement'…
7. …' manufactur*' …'by' …'plaintiff' …
8. …'not"… 'author*'…
9. …'defend*'… 'unauthorize'…
10. …'continu*'… 'us*'… 'plaintiff'… 'trademark'…
11. …'continu*'… 'infring*'… 'plaintiff'… 'trademark'…
12. …'maintain*'… 'us*'… 'plaintiff'… 'trademark'…
13. … 'maintain*'… 'infring*'… 'plaintiff'… 'trademark'…

be extracted using the related terms and key phrases that are of significance to the trademark case under review.

4. CASE IMPLEMENTATION

The case study uses python text mining to automatically extract the elements and precedent factors relevant to the recommender system query. One hundred and fifty judicial precedents related to trademark infringement, retrieved from FindLaw.com, are selected as the research target. The common and independent factors are extracted by the computer to identify issues and litigation judgment trends. The databased logical inferences are used to analyze the correlation between agreement status, applicable law, and judgment result. Finally, the TM infringement litigation evolution assists TM owners to protect their intangible assets and identify legal precedents that may be strategically used by their legal team for proactive brand defense.

4.1. Data Collection of TM Infringement Precedents

FindLaw is a free legal information website, launched in 1995 and is owned by Thomson Reuters Corporation (FindLaw, 2018). The website helps consumers, small-business owners, students and legal experts obtain relevant legal resources, related news, and legal analysis (FindLaw, 2018). This research extracts 150 judicial precedents from FindLaw for Legal Professionals (https://lp.findlaw.com/) using the key words "trademark infringement" and "initial interest confusion" during the years 2000 to 2015.

4.2. Extracting Precedent Key Factors Using Text Mining

This study uses text mining to retrieve eleven essential factors from the precedents which are formatted as abstracts for concise knowledge representation. The concise form enables TM owners to quickly clarify the independent factors of the legal precedent such as the agreement status, infringement type, applicable law, and judgment result. In Table 3, the independent factors are used as the key to find similar precedents for protecting the trademark. For instance, the contents of four independent factors are regarded as high relevance key terms for extracting three similar precedents based on the

Table 3. Examples for precedent recommendations

Precedent Name	Independent Factor			
	Agreement Status	Applicable Law	Judgment Result	Infringement Type
(Target precedent) Brookfield Communications v. West Coast Entertainment Corporation	No agreement	15 USC§1057 15 USC§1065 15 USC§1114 15 USC§1125 15 USC§1127	Remand	Similar trademark
				1. Strength or weakness of the plaintiff's mark 2. Marketing/advertising channels. 3. Product line expansion
(Similar precedent A) Sensient Technologies Corporation v. Sensoryeffects Flavor Company	No agreement	15 USC§1114 15 USC§1125 15 USC§1127	Remand	Similar trademark
				1. Strength or weakness of the plaintiff's mark 2. Similarity of plaintiff's and defendant's mark 3. Actual confusion 4. Defendant's intent 5. Marketing/advertising channels. 6. Product line expansion
(Similar precedent B) Multi Time Machine v. Amazon.com	No agreement	15 USC§1114 15 USC§1125 28 USC§1291	Conviction	Same trademark
				1. Strength or weakness of the plaintiff's mark 2. Actual confusion 3. Consumer's degree of care 4. Marketing/advertising channels.
(Similar precedent C) McNeil Nutritionals v. Heartland Sweeteners	No agreement	15 USC§1121 15 USC§1125 28 USC§1292 28 USC§1338	Remand	Similar trademark
				1. Strength or weakness of the plaintiff's mark 2. Defendant's use of the mark 3. Similarity of plaintiff's and defendant's mark 4. Actual confusion 5. Defendant's intent 6. Marketing/advertising channels. 7. Consumer's degree of care 8. Product line expansion

target precedent. The strength or weakness of the plaintiff's mark, marketing, and advertising channel distribution are the most common infringement violations.

4.3. Correlation Between Independent Factors

The correlation between agreement status, applicable law, and judgment result are of vital importance. The agreement status of a precedent influences the applicable law and the subsequent judgment. This research integrates the agreement status of 150 precedents, the frequency of applicable laws, and the correlation between judgments.

4.3.1. Agreement Status

In Table 4, the study integrates the agreement status of 150 randomly selected precedents which are divided into three periods (i.e., 2000-2005, 2006-2010, 2011-2015). Each proportion of no agreement between the plaintiff and defendant is higher than agreement.

4.3.2. Applicable Law

Text mining technique extracts all quoted laws and highlights the frequency applicable to the 150 precedents. The top three applicable laws are 15 USC§1114, 15 USC§1125, and 15 USC§ 1127. Table 5 summarizes the scope of the USPTO's laws and regulations. According to the scope of these laws, infringement, publication, inappropriate description, and the intention of the infringer are the most frequently disputed. Legal citation 15 USC§1125 occurs for almost half the sample cases (48.7% of the 150 precedents).

4.3.3. Judgment Result

Regarding the 150 precedents, 42 precedents are judged as acquittal, 47 precedents are judged as conviction, and 61 precedents are judged as remand. The correlation between judgment results and usage laws are shown in Figure 11. For remand, the usage frequency of the three common application laws are nearly equal. Referring to acquittal and conviction, the majority of the application laws quote 15 USC§1114 and 15 USC§1125. The analysis shows that over time, judgments are based on 15 USC§1114 and 15 USC§1125. The application of these two laws by judge's is reasonable. For example, law 15 USC1127 easily causes different judgments because it involves the intention of the infringer which falls under the judge's discretion.

4.4. The Evolution of TM Infringement Litigation

The eight infringement types are categorized as judgment results and are statistically depicted in Figures 12, 13, and 14. The figures show that earlier lawsuits (from 2000 to 2005) focus on the second to fifth infringement types (i.e., defendant's use of mark, similarity of plaintiff's and defendant's trademark, actual confusion, and the defendant's intent). During the period from 2006 to 2010, the second and fourth types (defendant's use of mark and actual confusion) are the most frequency used. The second, fourth, and fifth types (defendant's use of mark, actual confusion, and the defendant's intent) were widely applied during the years 2011 to 2015. Although the reasons behind these trends

Table 4. Agreement status of 150 precedents

	Have Agreement	**No Agreement**	**Total Precedents**
2000-2005	8 (16%)	42 (84%)	50
2006-2010	9 (18%)	41 (82%)	50
2011-2015	12 (24%)	38 (76%)	50

Table 5. The frequency laws applied (USPTO, 2017)

Law	**The Scope of the Law**
15 USC§1114	Remedies; infringement; innocent infringement by printers and publishers
15 USC§1125	False designations of origin, false descriptions, and dilution forbidden
15 USC§1127	Construction and definitions; intent of chapter

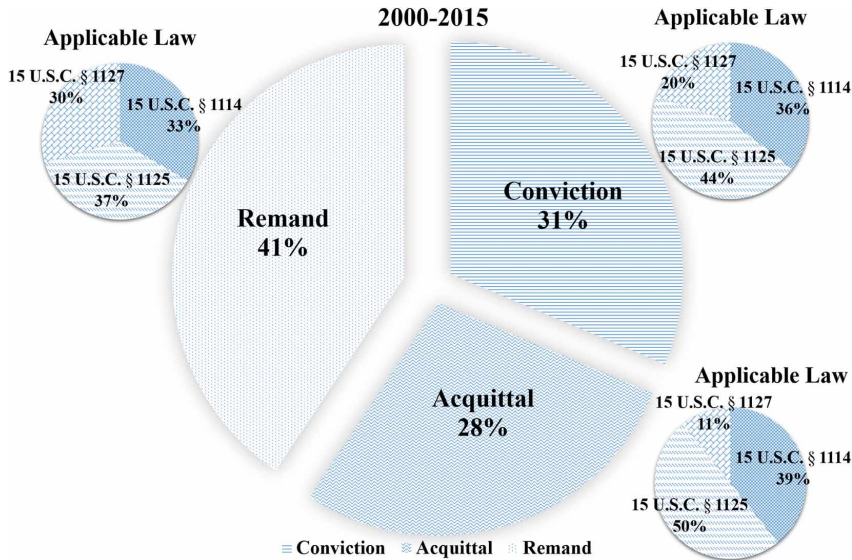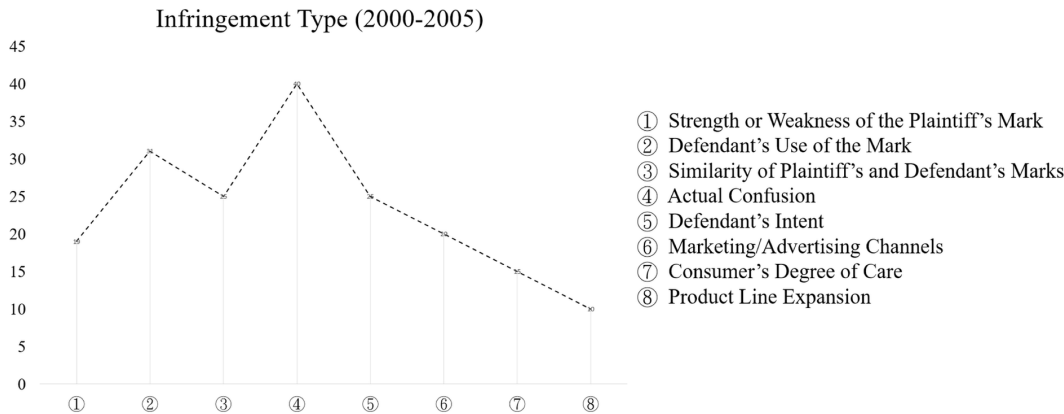**Figure 11. The correlation between judgment results and usage laws**

Figure 12. The statistics of trademark infringement (2000-2005)

① Strength or Weakness of the Plaintiff's Mark
② Defendant's Use of the Mark
③ Similarity of Plaintiff's and Defendant's Marks
④ Actual Confusion
⑤ Defendant's Intent
⑥ Marketing/Advertising Channels
⑦ Consumer's Degree of Care
⑧ Product Line Expansion

are beyond the scope of the present research, further investigation of the cases would help illustrate environmental changes in the marketplace and legal analysis.

The research combines the evolution of trademark litigation by combining the three curves in Figure 15. The results illustrate that the litigation of TM holders focus on the similarity of marks, actual confusion, and the marketing channels instead of the defendant's use of the mark. Further research is required to provide a reasonable explanation for these trends. One can only venture to say that internet commerce may be a good starting point for further research and analysis as a potential source of influence and change.

5. CONCLUSION

Corporate wealth is driven by IP rights as well as by the sale of tangible goods. According to the WIPO report (2017), more than one third of the value of manufactured products sold around the

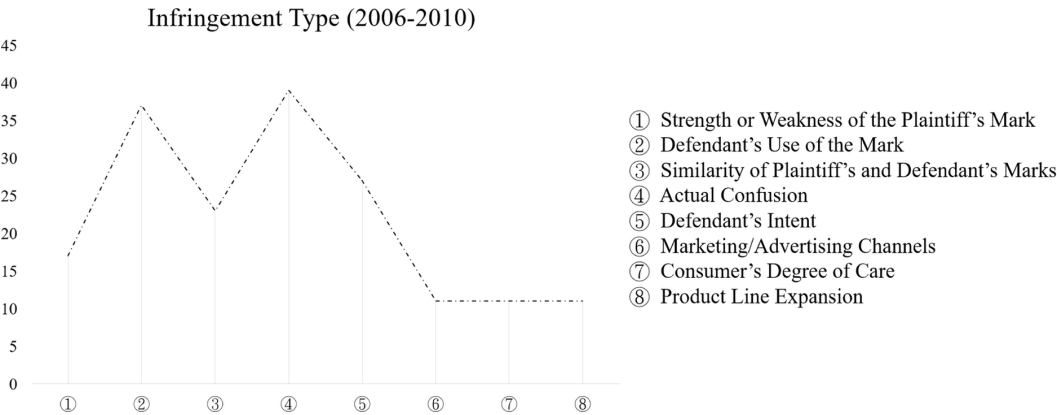**Figure 13. The statistics of trademark infringement (2006-2010)**

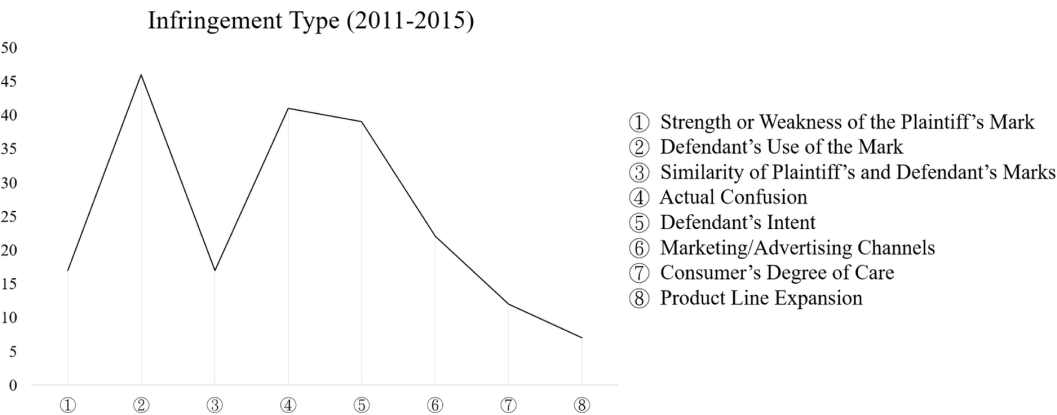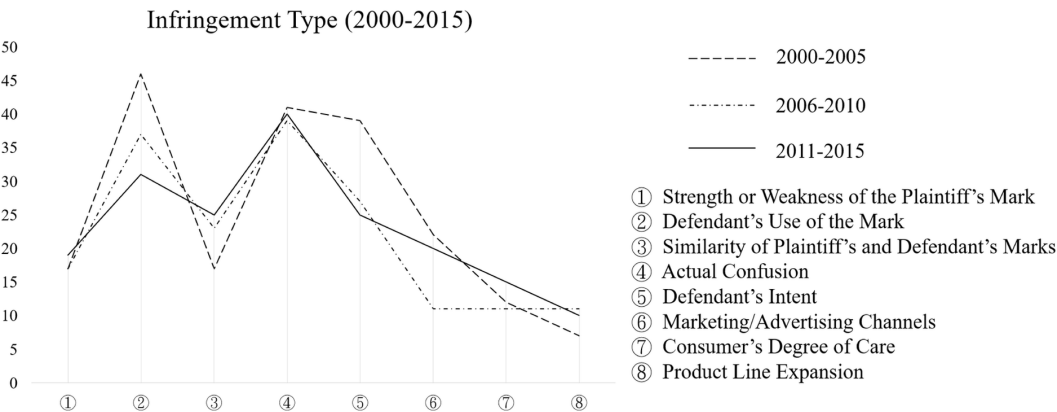**Figure 14. The statistics of trademark infringement (2011-2015)**

Figure 15. The evolution of trademark litigation (2000-2015)

world comes from intangible capital (e.g., branding, design, and technology). This is twice the value of tangible capital (e.g., buildings and machinery), emphasizing the growing role of IP in the world's economy. Clearly, intangible assets are vital to global value chains. If a business wishes to thrive globally, it needs to identify, protect and export its IP. As the trade in imitated products continues to grow online, anti-counterfeiting efforts and increased monitoring is needed. If companies are unable to stop online infringers, then new methods and means of analysis must be designed and applied.

In order to solve the challenge of monitoring potential trademark infringement and understand the logic that may be applied to resolve disputes, this research provides an intelligent trademark legal precedent recommender system. Trademark owners and lawyers are better enabled to quickly find relevant precedents, review the judgments, and develop strategies that will facilitate timely resolution of disputes. The text mining methods reduce manual costs and yields information more closely associated with the intangible assets that the company has a duty to protect. Trademark owners, legal counselors, IP experts, and other stakeholders can use the system to quickly find trademark precedent documents and study related judgments.

Our research provides an initial ontology and new techniques to protect trademarks and avoid infringing on the rights of TM owners. This research facilitates global strategic management and technology practices by exploring related trademark judicial precedents and their judgments, provides new analytic information about TM litigation, and provides a means to compare precedent data with potential infringement cases. The ontology depicts the knowledge structure of trademark infringement and helps to visualize trademark infringement precedents and understand the relationships between common and independent factors of the precedents.

The research demonstrates the development of an intelligent precedent recommender system using computer-aided drawing of a trademark ontology for a selected number of US judicial precedents covering the 2000-2015 year period. The management of intellectual property requires new and effective algorithms given the sheer volume of products that are being marketed via multi-channel and global digital networks. Image search and data clustering tools, as well as technology to capture and track infringing information is already required to defend IP. Future research will require the application of better algorithms to fully automate the processes and update the ontology and database in real time. Potentially, AI techniques will assist users to identify potential infringement problems and develop strategies to better protect trademarks across globally distributed electronic sales networks.

ACKNOWLEDGMENT

REFERENCES

Albert, G. P., Laff, W., & Laff, S. (1999). *Intellectual Property Law in Cyberspace*. BNA Books.

Allen, M. J. (1991). Who must be confused and when? The scope of confusion actionable under trademark law. *Trademark Reporter*, *81*(2), 209–259.

Berry, M. W. (2004). Survey of text mining. *Computer Review*, *45*(9), 548.

Cavazo, R. (2019). The economic cost of bad actors on the internet: Ad fraud. *Cheq*. Retrieved July 20, 2019, from https://www.mmaglobal.com/files/casestudies/cheq_the_economic_cost_of_bad_actors_on_the_internet.pdf

Cohen, D. (1995). *Legal Issues in Marketing Decision Making*. Cincinnati: South-Western Publishing.

Enberg, J. (2019, May 28). Global digital Ad spending 2019 – digital accounts for half of total media Ad spending worldwide. *eMarketer*. Retrieved July 20, 2019, from https://www.emarketer.com/content/global-digital-ad-spending-2019

Ertekin, L., Sorescu, A., & Houston, M. B. (2018). Hands off my brand! The financial consequences of protecting brands through trademark infringement lawsuits. *Journal of Marketing*, *82*(5), 45–65. doi:10.1509/jm.17.0328

Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the power of text mining. *Communications of the ACM*, *49*(9), 76–82. doi:10.1145/1151030.1151032

FindLaw Corporate Information. (2018). Retrieved July 19, 2019, from https://company.findlaw.com/company-history/findlaw-corporate-information-press-company-background.html

Hamidi, H., & Mousavi, R. (2018). Analysis and evaluation of a framework for sampling database in recommenders. *Journal of Global Information Management*, *26*(1), 41–57. doi:10.4018/JGIM.2018010103

Han, J., Huang, Y., Kumar, K., & Bhattacharya, S. (2018). Time-varying dynamic topic model: A better tool for mining microblogs at a global level. *Journal of Global Information Management*, *26*(1), 104–119. doi:10.4018/JGIM.2018010106

Hearst, M. A. (1999). Untangling text data mining. *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*, 3-10. doi:10.3115/1034678.1034679

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *Management Information Systems Quarterly*, *28*(1), 75–105. doi:10.2307/25148625

Howard, D. J., Kerin, R. A., & Gengler, C. (2000). The effects of brand name similarity on brand source confusion: Implications for trademark infringement. *Journal of Public Policy & Marketing*, *19*(2), 250–264. doi:10.1509/jppm.19.2.250.17131

Keller, K. L. (1998). *Strategic Brand Management: Building, Measuring, And Managing Brand Eqeuity*. Prentice Hall.

Lanham Act of. 1946 (2000). Chap. 540, 60 Stat. 427 (Codified as amended at 15 USC §1051–1157).

McCarthy, T. (1998). *McCarthy on Trademarks and Unfair Competition* (4th ed.). Deerfield, IL: Clark Boardman Callaghan (Suppl. 2005).

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, *24*(3), 45–78. doi:10.2753/MIS0742-1222240302

Raguseo, E., Pigni, F., & Piccoli, G. (2018). Conceptualization, operationalization, and validation of the digital data stream readiness index. *Journal of Global Information Management*, *26*(4), 92–112. doi:10.4018/JGIM.2018100106

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, *24*(5), 513–523. doi:10.1016/0306-4573(88)90021-0

Scott, C. D. (2013). Trademark strategy in the internet age: Customer hijacking and the doctrine of initial interest confusion. *Journal of Retailing*, *89*(2), 176–189. doi:10.1016/j.jretai.2012.11.004

Simoudis, E. (1996). Reality check for data mining. *IEEE Expert: Intelligent Systems and Their Applications*, *11*(5), 26–33. doi:10.1109/64.539014

Slefo, G. (2015, October 22). For every US 3 dollars spent on digital ads, fraud takes US 1 dollar. *AdAge*. Retrieved July 20, 2019, from https://adage.com/article/digital/ad-fraud-eating-digital-advertising-revenue/301017/

Staab, S., & Studer, R. (Eds.). (2004). *Handbook on Ontologies. International Handbooks on Information Sytems*. Springer.

Statista. (2018). *Super bowl average costs of a 30-second TV advertisement from 2002 to 2018*. Retrieved July 19, 2019, from https://www.statista.com/statistics/217134/total-advertisement-revenue-of-super-bowls/

Swartout, W., & Tate, A. (1999). Guest Editors' Introduction: Ontologies. *IEEE Intelligent Systems*, *14*(1), 18–19. doi:10.1109/MIS.1999.747901

Tan, A. H. (1999). Text mining: the state of the art and the challenges. *Proceedings of the PAKDD 1999 Workshop on Knowledge Discovery from Advanced Databases*, 65-70.

Trappey, C. V., Trappey, A. J. C., & Wu, C.-Y. (2010). Clustering patents using non-exhaustive overlaps. *Journal of Systems Science and Systems Engineering*, *19*(2), 162–181. doi:10.1007/s11518-010-5134-x

Trappey, A. J. C., Trappey, C. V., & Hsu, P. T. (2016a). Patent portfolio analysis of e-payment services using technical ontology roadmaps. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2016)* (pp. 4824-4829). IEEE.

Trappey, A. J. C., Trappey, C. V., Chang, A. C., & Chen, L. W. L. (2016b). Using web mining and perceptual mapping to support customer-oriented product positions and designs. *Proceedings of 2016 ISPE International Conference on Transdisciplinary Engineering (TE 2016)*.

U. S. Patent & Trademark Office (USPTO). (2018). *Protecting Your Trademark: Enhancing your rights through federal registration*. Retrieved July 20, 2019, from https://www.uspto.gov/sites/default/files/documents/BasicFacts.pdf

U. S. Patent & Trademark Office (USPTO). (2017). *Resources and glossary*. Retrieved July 19, 2019, from https://www.uspto.gov/trademark/resources-and-glossary

U. S. Patent & Trademark Office (USPTO). (2017). *U. S. trademark law*. Retrieved July 19, 2019, from https://www.uspto.gov/sites/default/files/documents/trademark_rules_statutes_2017-7-21.pdf

WebProtégé Administrator's Guide. (2017). Retrieved July 20, 2019, from: https://protegewiki.stanford.edu/wiki/WebProtegeAdminGuide

World Intellectual Property Organization (WIPO). (2017). *World intellectual property report 2017 - Intangible capital in global value chains (Chapter1, pp. 21-41)*. Retrieved July 20, 2019, from https://www.wipo.int/edocs/pubdocs/en/wipo_pub_944_2017.pdf

World Intellectual Property Organization (WIPO). (1996). *Implications of the TRIPS Agreement on Treaties Administered by WIPO*. Retrieved July 20, 2019, from https://www.wipo.int/edocs/pubdocs/en/intproperty/464/wipo_pub_464.pdf

World Trade Organization (WTO). (1994). *Agreement on trade-related aspects of intellectual property rights (TRIPS Agreement)*. Retrieved July 20, 2019, from https://www.wipo.int/wipolex/en/other_treaties/text.jsp?file_id=305736

Yuce, S. T., Agarwal, N., Wigand, R. T., Lim, M., & Robinson, R. S. (2014). Bridging women rights networks: Analyzing interconnected online collective actions. *Journal of Global Information Management*, *22*(4), 1–20. doi:10.4018/jgim.2014100101

Charles V. Trappey received his PhD in Consumer Science and Retailing from Purdue (USA), an MS in Quantitative Business Analysis from Louisiana State University (USA), and an LLM in IP Law from Queensland University of Technology (Australia). Prof. Trappey's research interests are in innovation and technology forecasting, IP informatics, retail business development, and technology marketing. Dr. Trappey is an Australia Trademark Attorney.

Ai-Che Chang received her PhD in Industrial and Business Management from National Taipei University of Technology, Taiwan. Her research interests are in technology innovation, digital marketing, and IP issues related to marketing. Dr. Chang has ten years teaching experience and working experience focusing on promoting technology innovation.

Amy J. C. Trappey received her PhD in Industrial Engineering from Purdue University. Prof. Trappey's research interests are in knowledge engineering and intelligent systems, particularly for IP and patent analytics, e-business and e-manufacturing. Prof. Trappey is currently serving as the Associate Editor for Advanced Engineering Informatics and IEEE Transactions on SMC: Systems. Dr. Trappey is an ASME and ISEAM Fellow and a senior member of IEEE.